

# UC Berkeley

## UC Berkeley Previously Published Works

**Title**

Are the Central Flores languages really typologically unusual?

**Permalink**

<https://escholarship.org/uc/item/5p49t2w2>

**Author**

Elias, Alexander

**Publication Date**

2020-01-14

Peer reviewed

# Are the Central Flores languages really typologically unusual?

Alexander Elias

January 13, 2020

## 1 Abstract

The isolating languages of Central Flores (Austronesian) are typologically distinct from their nearby relatives. They have no bound morphology, as well elaborate numeral classifier systems, and quinary-decimal numeral system. McWhorter (2019) proposes that their isolating typology is due to imperfect adult language acquisition of a language of Sulawesi, brought to Flores by settlers from Sulawesi in the relatively recent past. I propose an alternative interpretation, which better accounts for the other typological features found in Central Flores: the Central Flores languages are isolating because they have a strong substrate influence from a now-extinct isolating language belonging to the Mekong-Mamberamo linguistic area (Gil 2015). This explanation better accounts for the typological profile of Central Flores and is a more plausible contact scenario.

Keywords: Central Flores languages, Eastern Indonesia, isolating languages, Mekong-Mamberamo linguistic area, substrate influence

## 2 Introduction

The Central Flores languages (Austronesian; Central Malayo-Polynesian) are a group of serialising SVO languages with obligatory numeral classifier systems spoken on the island of Flores, one of the Lesser Sunda Islands in the east of Indonesia. These languages, which are almost completely lacking in bound morphology, include Lio, Ende, Nage, Keo, Ngadha and Rongga. Taken in their local context, this typological profile is unusual: other Austronesian languages of eastern Indonesia generally have some bound morphology and non-obligatory numeral classifier systems. However, in a broader view, the Central Flores languages are typologically similar to many of the isolating languages of Mainland Southeast Asia and Western New Guinea, many of which are also isolating, serialising SVO languages with obligatory numeral classifier systems.

The question of how the Central Flores languages became isolating is addressed by McWhorter (2019): he claims that Central Flores languages lost their morphology because they were acquired by a large number of adult speakers in the relatively recent past, perhaps arriving from Sulawesi. Under this account, the Central Flores languages lost their morphology due to imperfect learning and simplification by adult learners, and are an instance of a more general process which is exemplified by creole languages.

However, an explanation based on simplification alone cannot account for the other typological parallels between Central Flores, Mainland Southeast Asia and Western New Guinea, particularly the presence of complex numeral classifier systems. Gil (2015) has proposed a Mekong-Mamberamo (MM) linguistic area spanning Mainland Southeast Asia, the Indonesian Archipelago and Western New Guinea, based on 17 shared typological features. The Austronesian family is a relative newcomer to the Mekong-Mamberamo linguistic area, and has displaced the genealogically diverse MM-type languages in most parts of the Indonesian Archipelago. However, these pre-Austronesian languages have left their mark on the modern Austronesian languages of Indonesia to varying degrees. If the Mekong-Mamberamo hypothesis is correct, it provides a more economical explanation for the typology of the Central Flores languages.

In this paper, I argue that McWhorter’s proposed scenario of relatively recent contact with Sulawesi is implausible, and the typology of Central Flores languages is better explained as a reflection of a substrate language with Mekong-Mamberamo typology. This substrate is stronger in Central Flores than in East and West Flores, reflecting differing contact conditions between the Austronesian settlers and the pre-Austronesian population at the time of the original Austronesian settlement of Flores between 2500-1500 BCE (Bellwood 1997).

The structure of this paper is as follows. The first section gives the theoretical background (section 3): I will introduce the Mekong-Mamberamo proposal which will guide the rest of the paper (section 3.1) and outline McWhorter’s view on isolating languages (section 3.2), McWhorter’s stance on language complexity (section 3.3) and his historical scenario for the development of the Central Flores languages (section 3.4). Section 4 is a typological overview of the languages of West Flores (section 4.1), East Flores (section 4.2) and an introduction to the languages of Central Flores (section 4.3). Section 5 makes up the bulk of the paper: in this section, I examine the list of Mekong-Mamberamo features and illustrate their presence or absence in the languages of Flores. I then introduce additional relevant data about the numeral systems of Central Flores (section 6) before offering my own historical interpretation of the data (section 7) and finishing with a conclusion (8).

### 3 Theoretical background

In this section, I will lay the theoretical groundwork needed to interpret the data presented in section 5. First, I briefly describe the Mekong-Mamberamo linguistic area proposal (put

forth in Gil 2015) and introduce the features which he identifies as typical of the area. These features will be defined and explained more fully in section 5 when I address their presence or absence in the languages of Central Flores. After introducing the Mekong-Mamberamo proposal, I outline the theoretical framework from which McWhorter (2019) approaches the question of Central Flores typology. I outline his thinking on how languages become simplified by adult language acquisition (section 3.2), the criteria he proposes to evaluate linguistic complexity (section 3.3) and his historical proposal for the Central Flores languages (section 3.4).

### 3.1 The Mekong-Mamberamo language area

Based on typological similarities between the languages of Mainland Southeast Asia, the Indonesian Archipelago and Western New Guinea, Gil (2015) proposes the existence of the Mekong-Mamberamo linguistic area defined by the following 17 typological features.

1	passing gesture
2	repeated dental clicks expressing amazement
3	conventionalised greeting with ‘where’
4	‘eye day’ → ‘sun’ lexicalisation
5	d/t place-of-articulation asymmetry
6	numeral classifiers
7	verby adjectives
8	basic SVO constituent order
9	iamitive perfects
10	‘give’ causatives
11	low differentiation of adnominal attributive constructions
12	weakly developed grammatical voice
13	isolating word structure
14	short words
15	low grammatical-morpheme density
16	optional thematic-role flagging
17	optional TAM marking

Table 1: List of Mekong-Mamberamo typological features (Gil 2015; p267)

Gil (2015) proposes that the typological similarities between Mekong-Mamberamo languages reflect an ancient pattern of cultural contact across the area leading to linguistic convergence. The Mekong-Mamberamo area is proposed to be of great antiquity, pre-dating the arrival of the Austronesians in the Indonesian Archipelago around 2500-1500 BCE (Bellwood 1997). When the Austronesians arrived in the Mekong-Mamberamo linguistic area, they brought with them a distinctly non-Mekong-Mamberamo type language: verb-initial, with copious morphology and a well-developed system of grammatical voice.

This typological profile describes many of the modern Austronesian languages of Taiwan and the Philippines, and is the typological profile inherited from Proto-Austronesian (Blust 2013).

However, the Austronesian speakers who went south to the Indonesian Archipelago and spread east and west encountered speakers of Mekong-Mamberamo type languages where they settled. Eventually, almost all pre-Austronesian Mekong-Mamberamo type languages of the Indonesian Archipelago were displaced by Austronesian languages, but in the process they were restructured to fit the Mekong-Mamberamo typological profile to varying degrees. The degree to which any particular Austronesian language was restructured must have depended on a number of factors, including the ratio of settlers to local population on an island, the intensity and nature of the contact between them, and the social relationships between the settlers and the locals.

Thus, the Mekong-Mamberamo typological features display a saddle-shaped geographical distribution in many cases: they are most common in Mainland Southeast Asia and Western New Guinea area, with a patchier distribution across the Indonesian Archipelago. This is because of the incomplete restructuring of many Austronesian languages, which displaced the pre-Austronesian languages without fully conforming to the Mekong-Mamberamo typological profile yet.

The argument which I will develop in this paper is that the Central Flores languages are an example of a particularly heavily restructured group of Austronesian languages which have conformed almost totally to the Mekong-Mamberamo profile. Thus, they appear typologically unusual relative to many other Austronesian languages, but appear typologically well-behaved when seen as part of the Mekong-Mamberamo linguistic area. The typological differences between Central Flores and other parts of Flores reflect differences in the circumstances of contact between the Austronesian settlers and the pre-existing population (see section 7). Various other lines of evidence, such as that from Central Flores numeral systems (see section 6) back up this scenario.

### **3.2 McWhorter’s view of isolating languages**

The account of Central Flores typology outlined above and which I will argue for in this paper stands in contrast to McWhorter (2019), who also seeks to account for the isolating typology of the Central Flores languages. He proposes that a group of settlers from Sulawesi, perhaps speaking a language ancestral to *Tukang Besi*, arrived in Flores in the relatively recent past (ie, once Flores was already inhabited by Austronesian speakers). Their language was adopted by the pre-existing Austronesian speakers of Central Flores and was acquired by large numbers of adults. The process of imperfect adult language acquisition at that time resulted in the loss of bound morphology as adult speakers simplified the grammar. Thus, in McWhorter’s view, the Central Flores languages emerged from the simplification of *Tukang Besi* or some other language of Sulawesi, driven by imperfect adult language acquisition. Central Flores isolating typology is the result of general cognitive

processes at play whenever a language is imposed on a group of adult speakers (with creole languages best exemplifying this process of simplification).

McWhorter has developed an argument in a series of publications (2001, 2007, 2008, 2011, 2016, 2019) that highly isolating languages do not come about in situations of unbroken language transmission, but must always be the product of an episode of intense contact leading to imperfect adult language acquisition. In McWhorter's words, 'isolating typology *signals* heavy adult acquisition in a language's past, rather than merely *suggesting* it.' (McWhorter 2019; p193, emphasis in original). McWhorter's claim is as follows: because there is no other diachronic mechanism by which languages achieve such a totally isolating morphosyntactic profile, all 'radically' isolating languages must logically have undergone an episode of intense contact in the past which led to their current typological profile.

This is related to the distinction between 'esoteric' and 'exoteric' languages drawn by Thurston (1987): 'esoteric' languages are used only for communication within a small and tightly-knit group, where adult acquisition of the language is rare, while 'exoteric' languages are used for intergroup communication and as such, are commonly learned by adults. The process of adult learning strips away user-unfriendly opacities such as suppletion, irregularities and complex morphophonological alternations, and leaves an 'exoteric' language with less overall complexity than its 'esoteric' sisters.

In keeping with this view, McWhorter claims that abundant affixal morphology is the 'natural state' of language when it is transmitted uninterrupted between generations, given the vast learning capacity of infants. As the argument goes, it is inevitable that irregularities and opacities will accrue in a language which is learned only by infants because they have no strong need to restore systematicity. On the other hand, adult learners will seek to extend regularities and reduce opacities, because their language learning capacity is severely limited compared to that of infants.

The implication of this line of argument is that isolating languages do not stay isolating for long under regular conditions of intergenerational transmission. This serves as a kind of linguistic timer: when faced with an isolating language, one must not only posit a contact event, but it must be of rather recent date.

In this paper, I seek to show instead that the predictive strength of McWhorter's hypothesis (ie, simple languages only ever arise due to imperfect adult language learning) leads him to propose an unsound historical scenario of recent contact with Sulawesi to account for the typology of the Central Flores languages. McWhorter's hypothesis predicts that any changes occurring as a result of imperfect adult language learning tend towards simplification as he defines it. In this case, it is difficult for McWhorter's explanation to account the development of an elaborate system of classifiers in the Central Flores languages, which is a form of overspecification.

### 3.3 McWhorter's definition of linguistic complexity

In order to formalise his argument about relative complexity of languages, McWhorter has attempted to measure linguistic complexity - a notoriously difficult task - along three axes (McWhorter 2007; p21-35):

1) Overspecification: 'Languages differ in the degree to which they overtly and obligatorily mark semantic distinctions' (McWhorter 2007; p21). A language is more complex to the extent that it requires overt marking of person and number, noun class, definiteness, evidentiality, clusivity, tense, aspect, mood, etc...

2) Structural Elaboration: 'An aspect of one grammar may differ from that aspect in another's in terms of the number of rules (in phonology and syntax) or foundational elements (in terms of phonemic inventory) required to generate surface forms' (McWhorter 2007; p 29). A language is more complex to the extent that there are more unpredictable morphophonemic alternations, a larger phonemic inventory, more inflectional classes, word order alternations, etc...

3) Irregularity: 'Grammars differ in the degree to which they are festooned with irregularity and suppletion' (McWhorter 2007; p33). A language is more complex to the extent that its noun class system has arbitrary assignment, various unpredictable plural marking strategies, suppletion in its conjugational system, etc...

The purpose of this paper is not to dispute the fact that the Central Flores languages are relatively simple by McWhorter's metric of complexity (for a refutation of McWhorter's claim that certain languages of Timor are unusually simple by his own metric, see Schapper, this volume). They do indeed stand out in their local context as unusually isolating and devoid of opacities and irregularities.

However, many scholars would dispute McWhorter's complexity criteria, and much ink has been spilled trying to argue for and against various interpretations of linguistic complexity. To take one example, Fenk-Oczlon and Fenk (2008; p56) point to the fact that creole languages, often with simple morphology and phonology, tend towards a high level of polysemy and homophony. The Central Flores languages certainly tend towards polysemy, such as between intransitive and transitive uses of verbs (see the end of section 5.7). Under McWhorter's definition that contributes to overall simplicity because valency changes are not overtly marked on verbs. However, Fenk-Oczlon and Fenk (2009) point out that this massively increases the semantic complexity of the language, since each polysemous lexical item must still be associated with the proper range of possible constructions somehow.

### 3.4 McWhorter’s proposed historical scenario

In keeping with his views on the origins and development of isolating languages, McWhorter seeks to explain the typology of Central Flores by reference to either 1) a relatively recent migration from Sulawesi to Flores by speakers of a language similar or ancestral to Tukang Besi, or 2) contact with *Homo floresiensis*, a species of small hominid recently described from a handful of skeletons found in a cave in northwestern Flores dated to around 12,000 years ago (Brown et al. 2004). This second option is rather fanciful, and even assuming that Austronesian speakers co-existed with *Homo floresiensis* at some point, this would require that the imperfectly acquired speech of *Homo floresiensis* then became the dominant language of the entire community, even as they were pushed to extinction by modern humans. In my opinion, it is safe to lay the *Homo floresiensis* idea to rest, but the first scenario deserves a more careful look.

In this context, ‘relatively recent contact’ means that the contact occurred well after the initial contact between the incoming Austronesians and the earlier non-Austronesian (‘Papuan’) population of Flores, which may be placed between 2500-1500 BCE (Bellwood 1997). The occurrence of this initial contact is uncontroversial, because Flores was certainly inhabited at the time of the Austronesian settlement. However, McWhorter believes that little or nothing can be recovered about the earlier non-Austronesian languages of Flores, and that any proposals about them will ultimately lead to a scientific dead end: ‘One might propose that the central Flores languages became isolating in contact with now-extinct Papuan languages that were also isolating. This is reasonable - but a scientific dead end... These hypothetical isolating Papuan languages of Flores could only remain, therefore, an unverifiable surmise, whereas this paper is an attempt to assign a more systematic and refutable explanation to the facts’ (McWhorter 2011; p252).

Thus, to explain the isolating typology of Central Flores, McWhorter departs from the consensus to propose a second episode of contact, where Austronesian speakers from Sulawesi migrated to Flores, then shifted to the local languages but left traces of their imperfect adult language acquisition in the isolating typology of the modern Central Flores languages. However, as McWhorter himself points out, ‘lexical and grammatical data in support of this scenario are lacking’ (McWhorter 2019; p195). Indeed, there is a conspicuous absence of parallels between the languages of Central Flores and Sulawesi in lexicon, grammar and phonology. This absence is all the more conspicuous because the newcomers from Sulawesi would almost certainly have been in a socially dominant position over the local population, given that they were economically and technologically advanced enough to launch an overseas expedition.

In addition, numerous rulers in the region have traditionally established their claim to legitimacy by reference to foreign origin, including the Sika-speaking kingdom of East Flores (cf. Lewis 2010 *The Stranger-kings of Sikka*). The theme of an immigrant ‘stranger-king’ or ‘xenarch’ arriving from overseas and establishing a dynasty is prevalent throughout the eastern Lesser Sunda Islands (Lewis 2010), so there is precedent for the notion that at



least some groups of outside settlers enjoyed a high level of prestige in the area. Following the predictions of Thomason and Kaufman (1988), we would expect to find many lexical traces of the dominant group - the situation would be analogous to the Norman conquest of England by socially dominant but numerically inferior French speakers, leading to the shift of French speakers to English but with heavy lexical influence of French on the resulting English language. The alternative, that the arrivals from Sulawesi became integrated as the equals or the inferiors of the local population, is less plausible.

The evidence adduced by McWhorter in favour of this contact with Sulawesi is rather circumstantial, drawn from history and folklore, and is hardly the smoking gun which allows us to draw a direct link between Sulawesi and Flores: ‘The Gowa empire of southwestern Sulawesi controlled the Manggarai region of Flores from 1658 to 1750, and many Manggarai trace their ancestry to migrations from Gowa on the southwestern leg of Sulawesi (Erb 1999: 85-86). One of the ancestor stories of the Nage involves invaders from Gowa as well (Forth 1998: 230) and their cosmology traces them in general to either Sulawesi or ‘Bugis bonerate’. Manggarai and Nage people also trace ancestry to what they term the Minangkabau (Erb 1999: 85, Forth 1998: 81) but Van Bakkum (1944) documented the alternate term ‘Bonengkabau’, suggesting that ‘Minangkabau’ may be a folk distortion of an actual descent from the more geographically plausible region of the Gulf of Bone between the southwestern and southeastern legs of Sulawesi’ (McWhorter 2019: 194).

A lexical line of evidence, originally put forth by Hull (1998) while proposing a migration from Sulawesi to Timor, is taken up in McWhorter (2019) and adapted to the Flores context. This argument states that the rate of cognate matches between Sulawesi and Flores is very high, and that the forms of the cognates are so similar that they cannot have been separated for 3000 years without a fresh injection of Sulawesi lexicon into the Flores languages. In other words, pairs of languages separated for that long should undergo more evolution from their common source than is actually observed. In an earlier work, he cites pairs from *Tukang Besi* and languages of Flores such as the following, reproduced as is from McWhorter (2011; p241) with a few minor errors in the Ende data corrected. The Proto-Malayo-Polynesian forms on the right have been added by me, drawn from the Austronesian Comparative Dictionary (Blust and Trussell 2019):

Gloss	PMP	T. Besi	Sika	Ende	Ngadha	Rongga	Keo
‘come’	*maRi	mai	mai	mai	mai	mai	maʔi
‘dead’	*matay	mate	mate	mata	mata	mata	mata
‘fish’	*hikan	ika	iaŋ	ʔika	ika	ika	ʔika
‘fowl’	*manuk	manu	manu	manu	manu	manu	manu
‘liver’	*qatay	ate	wateŋ	ʔate	ate	ate	ʔate
‘pig’	*babuy	wawu	wawi	wawi	wawi	wawi	wawi
‘rain’	*quzan	usa-	uran	ʔura	uza	nura	ʔura
‘stone’	*batu	watu	watu	watu	vatu	watu	watu

Table 2: Lexical similarities between Tukang Besi and Flores languages are retentions, not innovations

This argument is flawed because it rests on shared retentions to support claims about historical relatedness. The forms cited by McWhorter are minimally changed from Proto-Malayo-Polynesian, and hundreds of additional examples of languages with similar forms could be adduced from across the Austronesian family. Shared retentions can never be taken as evidence in orthodox comparative linguistics; conservative languages are similar because of their relation to their common ancestor, not to each other. In order for this evidence to support a link between Tukang Besi and Flores, it would be necessary for McWhorter to show that the lexical similarities are in fact innovations, which they are not.

Another argument which McWhorter marshals in support of a recent contact hypothesis is the clinality of isolating languages in Flores. He states that ‘the languages become increasingly less isolating westward and eastward’. He suggests that this shows that there was total loss of affixation at one place (the landing site of the invaders from Sulawesi, presumably) which radiated outwards to neighbouring languages with less and less intensity. This is in fact not the case: the isolating languages of Central Flores form a well-defined clade, and there is ample evidence that the isolating profile of these languages can be reconstructed to their common ancestor, Proto-Central Flores (Elias 2018). Within the Central Flores-speaking area, isolating morphology is the rule, but the borders of this area to the west and east are sharp, not a gradient as McWhorter suggests. Although it is true that Sika is somewhat less complex than Lamaholot, it still retains a system of verbal conjugation and other morphological complexities which put it in a separate class from the Central Flores languages typologically. The pattern indicates that the current distribution of isolating morphology in Flores is not because of diffusion through contact, but rather because of common descent from a single, highly isolating Proto-Central Flores ancestor. This shows up very clearly in the linguistic data as an easily reconstructible node at Proto-Central Flores with well-defined bundles of isoglosses delimiting the boundaries of Central Flores (see section 4.3 for a list of innovations).

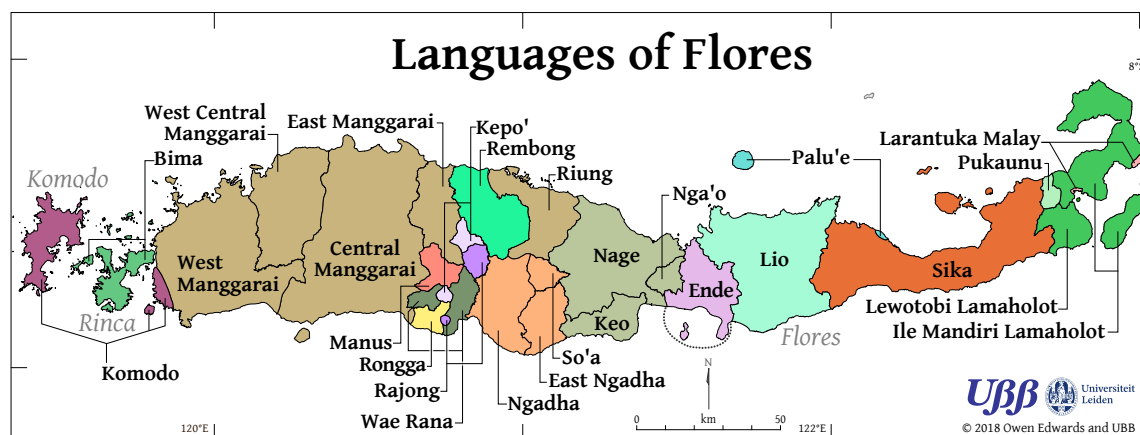
## 4 Introduction to the languages of Flores

The island of Flores is part of the Lesser Sunda Islands chain, located in the east of Indonesia (Nusa Tenggara Timur Province). Flores has a population of nearly two million as of the 2010 Indonesian census, and these people all speak Austronesian languages of the Central Malayo-Polynesian (CMP) group.

Linguistically, Flores can be divided into three approximately equal sections: West, East and Central Flores. West Flores is dominated by Manggarai (with a few poorly known languages similar to Manggarai spoken on the peripheries) while East Flores is populated by speakers of Sika and Lamaholot. Across Central Flores stretches the Central Flores Linkage: Lio, Ende, Nage, Keo, Ngadha and Rongga.

Blust (2008) finds some evidence the languages of Central Flores subgroup with West Flores (Manggarai) and the languages of nearby Sumba and Hawu, in a primary branch of CMP dubbed ‘Flores-Sumba-Hawu’, while East Flores (Sika, Lamaholot on East Flores plus Kedang, Alorise on neighbouring islands) belongs to a separate branch of CMP dubbed ‘Flores-Lembata’. Fricke (2019; p229) presents evidence that Flores-Sumba-Hawu and Flores-Lembata form a higher-order subgroup along with Bima, a group called ‘Bima-Lembata’.

Figure 1: Map of the languages of Flores (created by Owen Edwards 2018, reproduced with permission)



### 4.1 Languages of West Flores: Manggarai

The western third of Flores is dominated by Manggarai, the largest CMP language by number of speakers. There are a few other poorly known languages in West Flores (Kepo', Rembong, Riung, Manus, Rajong, Wae Rana) which appear to be close to Manggarai

based on lexical inspection but remain largely uninvestigated. There is also the better-known Komodo language spoken in the very west of Flores and on neighbouring Komodo Island. For the purposes of this paper, Manggarai will represent all of West Flores, because Komodo is known to be very similar, and the other languages are too scantily known to draw conclusions from. The bulk of the work on these languages has been carried out by Verheijen, a Dutch linguist, including his monumental Manggarai dictionary (Verheijen 1967) and sketch grammar of Komodo (Verheijen 1982).

The Central Manggarai variety described by Semiun (2017) has a modest amount of morphology. It has a set of enclitics which cross-reference the subject of the verb, shown in Table 3. These enclitics need not attach to the verb, but typically attach to the last word of the clause, hence their status as clitics.

Person	Form	Gloss
1SG	aku haŋ=k	‘I eat’
2SG	hau haŋ=h	‘You eat’
3SG	hia haŋ=i	‘He/She eats’
1PL.INC	ite haŋ=t	‘We (incl.) eat’
1PL.EXC	ami haŋ=km	‘We (excl.) eat’
2PL	meu haŋ=m	‘You (pl.) eat’
3PL	ise haŋ=s	‘They eat’

Table 3: Manggarai subject-marking clitics

There is a second set of possessive enclitics indicating the possessor, and which differ in some cases from the subject-marking enclitics, as shown in Table 4.

Person	Form	Gloss
1SG	mbaru=k	‘my house’
2SG	mbaru=m	‘your house’
3SG	mbaru=n	‘his/her house’
1PL.INC	mbaru=t	‘our (incl.) house’
1PL.EXC	mbaru=km	‘our (excl.) house’
2PL	mbaru=s	‘your (pl.) house’
3PL	mbaru=d	‘their house’

Table 4: Manggarai possessive clitics

Manggarai uses a decimal number system, with familiar Austronesian numerals: *tʃa* ‘one’, *sua* ‘two’, *təlu* ‘three’, *pat* ‘four’, *lima* ‘five’, *ənəm* ‘six’, *pitu* ‘seven’, *alo* ‘eight’, *tʃiok* ‘nine’, *pulu* ‘ten’ (Verheijen 1967).

Manggarai has numeral classifiers, but these are not obligatory when using numerals:

- (1) *pulu wuykut*  
 ten knuckle  
 ‘ten knuckles’ (Manggarai: Verheijen 1967; p8)

## 4.2 Languages of East Flores: Sika, Lamaholot

East Flores has two languages, Sika and Lamaholot, which belong together in the Flores-Lembata subgroup of CMP. These languages show a significant amount of morphological complexity, both inflectional and derivational.

The data presented on Sika is drawn from Fricke (2013), which describes aspects of the grammar of Hewa, an eastern variety of Sika. There is also a Sika dictionary available (Pareira and Lewis 1998).

A subset of Sika verbs is conjugated with an initial consonant cross-referencing the subject (see Table 5). The full set of conjugations appears only with vowel-initial verb roots. In consonant-initial verb roots, the initial consonant undergoes alternations which are mostly predictable from the voicing of the subject marker found in vowel-initial words. There is one irregular verb /ʔa/ ‘to eat’. This is not the full extent of verbal conjugation in Sika, but it gives a flavour of the type of alternations encountered.

Person	Pronoun	/-inu/ ‘to drink’	/pano/ ‘to go’	/ʔali/ ‘to dig’	/ʔa/ ‘to eat’
1SG	aʔu	ʔ-inu	pano	ʔali	ʔoa
2SG	ʔau	m-inu	bano	gali	goa
3SG	nimu	n-inu	bano	gali	ga
1PL.INC	ʔita	t-inu	pano	ʔali	ʔea
1PL.EXC	ʔami	m-inu	bano	gali	gea
2PL	miu	m-inu	bano	gali	gea
3PL	rimu	r-inu	pano	ʔali	ʔa

Table 5: Sika (Hewa) verb conjugation

In possessive constructions, a morpheme /-n/ is added to the second member of the construction. This can be the possessor, as in the pronominal possessive construction [Noun + Pronoun]:

- (2) *me nimu-n*  
 child 3SG-POSS  
 ‘his/her child’ (Sika (Hewa): Fricke 2013; p39)

The same possessive morpheme /-n/ can appear on the possessum, in the nominal possessive construction [Possessor + Noun]:

- (3) *duʔa ʔia me-n*  
 woman DEM child-POSS  
 ‘that woman’s child’ (Sika (Hewa): Fricke 2013; p40)

Sika has a decimal numeral system, with mainly inherited Austronesian numerals: *ha* ‘one’, *rua* ‘two’, *təlu* ‘three’, *hutu* ‘four’, *lima* ‘five’, *əna* ‘six’, *pitu* ‘seven’, *walu* ‘eight’, *hiwa* ‘nine’, *pulu* ‘ten’. The only numeral here which is not an inherited Austronesian form is *hutu* ‘four’, which is likely a loan from the Lio *sutu* ‘four’ with a regular change of /s/ to /h/ (Fricke 2019; p367).

Sika has numeral classifiers, but these are optional when using numerals. Thus, in the Sika construction *ʔita rua-t* ‘the two of us’, no numeral classifier is needed.

- (4) *ʔita rua-t*  
 1PL.INCL two-ATT  
 ‘the two of us’ (Sika (Hewa): Fricke 2013; p47)

The equivalent construction in a Central Flores language would be ungrammatical without a numeral classifier, as in the following Lio example:

- (5) *\*kita rua*  
 1PL.INCL two  
 Failed reading: ‘\*the two of us’ (Lio: Ibu\_Ferdy\_Frogstory)

- (6) *kita imu rua*  
 1PL.INCL CLASS.HUM two  
 ‘the two of us’ (Lio: Ibu\_Ferdy\_Frogstory)

The other language of East Flores, Lamaholot, shows considerable internal diversity and is also spoken in the neighbouring islands of Solor, Adonara and Lembata. A number of dialects of Lamaholot, both on Flores and on neighbouring islands, have received significant linguistic attention. There is a dictionary of the Lewolema dialect (Pampus 1999), a PhD thesis describing with the Lewotobi dialect (Nagaya 2011), a PhD thesis describing the Central Lembata dialect (Fricke 2019), a description of the morphology of the Lamalera dialect (Keraf 1978), a grammar of the Lewoingu dialect (Nishiyama and Kelen 2007), and a sketch grammar of the Solor dialect (Arndt 1937). The data presented in this section are drawn from Nagaya (2011) on the Lewotobi dialect, spoken in East Flores.

Lamaholot is the most morphologically complex language of Flores, with subject-marking prefixes on verbs, subject-marking enclitics, and possessive marking, among other morphology. Lewotobi Lamaholot uses subject-marking enclitics, as shown in Table 6. These do

not necessarily attach to the verb, hence their status as a clitic; but this need not concern us here.

Person	Form	Gloss
1SG	go lega=əʔ	‘I walk’
2SG	mo lega=ko	‘You walk’
3SG	na lega=aʔ	‘S/he walks’
1PL.INC	tite lega=kə	‘We (incl.) walk’
1PL.EXC	kame lega=kə	‘We (excl.) walk’
2PL	mio lega=kə	‘You (pl.) walk’
3PL	ra lega=ka	‘They walk’

Table 6: Lewotobi Lamaholot subject-marking clitics

In addition to the subject-marking enclitics, there is a subset of vowel-initial verbs which take subject-marking prefixes. These often redundantly mark the subject in conjunction with the subject-marking enclitics (which here show an /n-/ initial form due to the preceding nasal vowel):

Person	Form	Gloss
1SG	go k-enũ=nəʔ	‘I drink’
2SG	mo m-enũ=no	‘You drink’
3SG	na n-enũ=naʔ	‘S/he drinks’
1PL.INC	tite t-enũ=nə	‘We (incl.) drink’
1PL.EXC	kame m-enũ=nə	‘We (excl.) drink’
2PL	mio m-enũ=nə	‘You (pl.) drink’
3PL	ra r-enũ=na	‘They drink’

Table 7: Lewotobi Lamaholot verb conjugation

Lewotobi Lamaholot makes an alienability distinction in possessive constructions. In Lewotobi Lamaholot, a morpheme /-N/ surfaces on the second member in an inalienable possessive construction, and is realised as nasalisation on the final vowel.

- (7) *ika leĩ*  
 ika lei-N  
 Ika foot-POSS  
 ‘Ika’s foot’ (Lamaholot (Lewotobi): Nagaya 2011; p33)

In an alienably possessed construction, the second member is marked by an enclitic /=kə/ instead:

- (8) *ika doi=kə̃*  
 Ika money=POSS  
 ‘Ika’s money’ (Lamaholot (Lewotobi): Nagaya 2011; p33)

In addition to this inflectional morphology, Lewotobi Lamaholot has a number of derivational affixes (not all of which are still productive) which make it significantly more morphologically complex than the other languages of Flores. For instance, there is a process of ‘nasal substitution’ by which a verb can be nominalised by replacing the initial consonant with a nasal (and possibly an initial consonant as well):

Verb	Gloss	Noun	Gloss
pətə	‘to cut’	mətə	‘cutting board’
bitu	‘to fish with rod’	mnitu	‘fishing rod’
dira	‘to use a fan’	mnira	‘fan’
giʔa	‘to scratch’	kniʔa	‘match’

Table 8: Nominalisation by nasal substitution in Lewotobi Lamaholot

The Lewotobi Lamaholot numeral system is mainly an inherited Austronesian decimal system: *toʔu* ‘one’, *rua* ‘two’, *təlo* ‘three’, *pa* ‘four’, *lema* ‘five’, *namu* ‘six’, *pito* ‘seven’, *buto* ‘eight’, *hiwa* ‘nine’, *pulo* ‘ten’. The most interesting numeral here is *buto* ‘eight’, which is plausibly related to Proto-Central Flores \*wutu ‘four’. A semantic and formal parallel is found in Lamaholot’s relative Kedang *butu rai* ‘eight’, where *rai* means ‘many’ (Fricke 2019; p367-368).

Numeral classifiers are not obligatory when using numerals in Lewotobi Lamaholot:

- (9) *gula rua*  
 candy two  
 ‘two pieces of candy’ (Lamaholot (Lewotobi): Nagaya 2011; p159)

### 4.3 Languages of Central Flores: Lio, Ende, Nage, Keo, Ngadha, Rongga

The closely related languages of Central Flores form a linkage across the central third of Flores and are very similar in their morphosyntactic structure. The differences between the modern Central Flores languages consist mainly of lexical differences and regular sound correspondences between phonemes. There is strong evidence that they form a clade, descending from Proto-Central Flores (Elias 2018). This evidence includes the loss of bound morphology and all coda consonants, the restructuring of the decimal numeral system into a mixed-base quinary-decimal system (see section 6), as well as a hefty amount of innovative basic vocabulary (PCF \*kobe ‘night’, \*mbeʔo ‘to know’, \*toro ‘red’, \*ndate



‘heavy’, \*kleu ‘betel nut’, \*lidu ‘sky’, \*koe ‘to dig’, \*longo ‘back’, \*pawe ‘good’, \*te?e ‘mat, \*we?e ‘near’) and a few shared semantic shifts (PMP \*beRŋi ‘night’ becomes ‘when?’ and PMP \*laku ‘civet’ becomes ‘dog’) (Elias 2018; p123-125).

The level of description of the Central Flores languages is uneven. Keo is described in a PhD thesis (Baird 2002) and Rongga has a grammar (Arka 2016) and dictionary (Arka et al. 2011). For Ngadha, there is a rather outdated grammar (Arndt 1933b) and dictionary (Arndt 1961). There is a dictionary of Lio (Arndt 1933a), as well as an undergraduate thesis on it (Levi 1978). Finally, there is an unpublished dictionary of Ende available only in electronic format (Aoki and Nakagawa 1993), as well as a description of Ende phonology (McDonnell 2009). Nage has not received much attention from linguists specifically, but Nage culture and folk classification has been the subject of numerous publications by the anthropologist Gregory Forth (Forth 1998, 2004, 2008, 2009, 2016).

There is a group of language varieties transitional between Ende and Keo referred to as Nga’o (shown on the language map, Figure 1) which may be divergent enough to be classified as a separate language, although it has been referred to as a dialect of Ende by previous researchers (Aoki and Nakagawa 1993).

Palu’e, spoken on an island of the same name off the northern coast of Flores, is very similar to the Central Flores languages and deserves a mention. It is described by a dictionary (Donohue 2003) and an analysis of the voice system (Donohue 2005). However, it falls outside of the scope of some of the key innovations that define the Central Flores clade: for instance, it retains a final /-n/ morpheme marking the genitive, so it has not undergone total loss of morphology and final consonants. It retains a typical Austronesian decimal numeral system as well, and has not developed the distinctive quinary-decimal numeral system of the Central Flores languages. Because it does not participate in these key Central Flores innovations, it falls outside of the Central Flores languages as I define them (Elias 2018) and is their closest external relative.

My own fieldwork has dealt mainly with Lio, and my MA thesis provides an analysis of Lio phonology, along with a comparative analysis of the Central Flores languages and reconstruction of aspects of Proto-Central Flores. (Elias 2018). Note that the name Lio (/lio/) is often erroneously spelled Li’o, implying the presence of a glottal stop (\*liʔo/). There is no glottal stop in the name of the language, and the confusion arises from the counterintuitive convention used in Arndt’s Lio dictionary (Arndt 1933a), where the *absence* of a glottal stop is indicated by an apostrophe.

The data presented in this paper to illustrate the typology of the Central Flores languages will be drawn from my own field data on Lio, as well as the other two Central Flores languages with full grammars: Keo (Baird 2002) and Rongga (Arka 2016). This provides a good geographic sampling of the Central Flores linkage (Lio is in the east, Rongga is in the west, Keo is in the middle). Given the varying level of documentation of the other languages, I will restrict myself to presenting data from these three languages. Given the close similarities between the Central Flores languages, I do not expect that additional data from Ende, Ngadha and Nage would substantially change the argument laid out here.

## 5 Central Flores languages have typical Mekong-Mamberamo typology

In this section, I will show that the Central Flores languages in my sample (Lio, Keo, Rongga) show nearly all of the Mekong-Mamberamo features proposed by Gil (2015).

### 5.1 The passing gesture

The passing gesture, used when a person needs to pass through someone’s personal space, is as follows: ‘while walking, the gesturer bends the top half of the body forward, and ... extends the right forearm forward, with the hand oriented vertically, palm facing inward, as though forging a path through an imaginary thicket’ (Gil 2015; p270).

I have observed Lio speakers employ the passing gesture when passing between two interlocutors in a conversation or passing through a crowded room.

I do not have information on whether speakers of Rongga or Keo use the passing gesture, so I will mark them with ‘?’ in the feature table below, but it is highly likely that they also employ this gesture.

	Lio	Keo	Rongga
The passing gesture:	+	?	?

### 5.2 Repeated dental clicks expressing amazement

In the Mekong-Mamberamo area, repeated dental clicks are used as a paralinguistic expression of amazement, usually with a positive affect. In contrast, most English speakers use repeated dental clicks to express disapproval, written as *tsk tsk* in the USA and *tut tut* in the UK.

Lio speakers do indeed use repeated dental clicks to express amazement, based on first-hand observation.

I do not have information on whether speaker of Rongga or Keo use repeated dental clicks to express amazement, so I will mark them with ‘?’ in the feature table, but it is highly likely that they also use them.

	Lio	Keo	Rongga
Repeated dental clicks expressing amazement:	+	?	?

### 5.3 Conventionalised greeting with ‘where’

In the Mekong-Mamberamo area, conventionalised greetings tend to be formed with the question word ‘where’, as in Indonesian *mau ke mana*, literally ‘Where are you going?’.

This is true for common conventionalised greetings both in Lio and in Keo:

- (10) *mbana əmba*  
go where  
'Where are you going? (Lio; Elias 2017 fieldnotes)'

- (11) *kau nuka ena ?emba*  
2SG go.up LOC where  
'Where are you going? (Keo: Baird 2002; p440)'

I do not have information on conventional greetings in Rongga, so I will mark it with '?' in the feature table.

	Lio	Keo	Rongga
Conventionalised greeting with 'where':	+	+	?

#### 5.4 'eye day' to 'sun' lexicalisation

The concept of 'sun' in the Mekong-Mamberamo area is often lexicalised as a collocation meaning something like 'eye of the day', as in Indonesian *mata hari* 'sun (lit. eye of day)'. This holds true in Lio, Rongga and Keo, and the form \*mata ləɖʒa 'sun' is reconstructible to Proto-Central Flores (Elias 2018):

- (12) \**mata ləɖʒa*  
eye day  
'sun' (Proto-Central Flores: Elias 2018)

- (13) *mata ləɖʒa*  
eye day  
'sun' (Lio: Elias 2017 fieldnotes)

- (14) *mata ləɾa*  
eye day  
'sun' (Rongga: Arka et al. 2011; p128)

- (15) *mata dəɾa*  
eye day

‘sun’ (Keo: Baird 2002; p559)

All three of these languages have ‘eye day’ to ‘sun’ lexicalisation, so I will mark all three of them with ‘+’ on the feature chart.

	Lio	Keo	Rongga
‘eye day’ to ‘sun’ lexicalisation:	+	+	+

### 5.5 d/t place-of-articulation asymmetry

An asymmetry in the place of articulation of the coronal stops /t/ and /d/ has been noted in many Mekong-Mamberamo languages: /t/ is dental (and more laminal), while /d/ is alveolar (and more apical).

This mismatch in place of articulation is present in Lio (Elias 2018):

- (16) /ata/  
 [ʔa.t<sup>h</sup>a]  
 ‘person’ (Lio: Elias 2018; p19)’

- (17) /ada/  
 [ʔa.da]  
 ‘tradition, custom’ (Lio: Elias 2017 fieldnotes)

There is no mention of a mismatch in place of articulation for Keo in Baird (2002), which lists both /t/ and /d/ as alveolar apical stops. There is also no mention of a mismatch for Rongga in Arka (2016) which lists both /t/ and /d/ as alveolar stops. I will mark them as ‘?’ in the table of features, because this is a rather low-level phonetic feature that could easily be omitted in a grammar, so I am not certain of its absence in Keo and Rongga.

	Lio	Keo	Rongga
d/t place-of-articulation mismatch:	+	?	?

### 5.6 Numeral classifiers

A feature of Mekong-Mamberamo languages is the presence of a system of numeral classifiers, such as those famous from the Mainland Southeast Asian languages. Numeral classifiers are independent morphemes which occur when numerals modify nouns in an NP, and the choice of classifier depends on the noun.

There is a large and obligatory set of numeral classifiers in the Central Flores languages. Any time a numeral is used, whether attributively or predicatively, a classifier is obligatory.

- (18) *saʔo əsa təlu*  
house CLASS.GEN three  
'three houses' (Lio: Ibu\_Ferdy\_History)
- (19) *\*saʔo təlu*  
house three  
Failed reading: '\*three houses' (Lio: Ibu\_Ferdy\_History)
- (20) *ana kai kolo sutu*  
child 3G CLASS.HUM four  
'She has four children. (lit. Her children are four)' (Lio: Elias 2017 fieldnotes)
- (21) *\*ana kai sutu*  
child 3G four  
Failed reading: '\*She has four children. (lit. Her children are four)' (Lio: Elias 2017 fieldnotes)

The general classifier is Lio, Rongga *əsa*, Keo *ʔəsa*. When counting humans, the appropriate classifier is Lio *kolo*, Rongga *mori*, Keo *ŋgaʔe*, and when counting animals it is Lio, Rongga *eko*, Keo *ʔeko*.

In addition to these three common classifiers, there are hundreds of other classifiers which sort by size, shape, texture, function, and many other categories. The classifier Lio, Keo, Rongga *puʔu* ('trunk') is used for large, cylindrical objects such as trees, while Lio, Keo, Rongga *toko* ('bone') is used for smaller cylindrical objects like tubers and sticks. Some classifiers are very abstract: Lio has a classifier *wuŋa*, used for things which can potentially be used as weapons (machetes, bows, spears, digging sticks), and there is a dedicated classifier in Lio, Keo *papa* ('side') for things which naturally come in pairs: spouses, legs, arms, left and right sides of an object.

While the singular forms with the proclitic *sa=* display the expected Austronesian order (numeral-classifier), higher numerals in the Central Flores languages show an inverted order (classifier-numeral).

- (22) *jata ria sa=eko*  
eagle large SG=CLASS.ANIMAL  
'one large eagle' (Lio: Ibu\_Ferdy\_Frogstory)

- (23) *laki kolo təlu*  
 chief CLASS.HUM three  
 ‘three chiefs’ (Lio: Ibu\_Ferdy\_History)
- (24) *uwi kadzu toko rua*  
 tuber wood CLASS.STICK two  
 ‘two cassavas’ (Lio: Positional\_Elicitation)
- (25) *saʔo ha=ʔəsa*  
 house SG=CLASS.GEN  
 ‘one house’ (Keo: Baird 2002; p182)
- (26) *dako ʔeko rua ena wəwa*  
 dog CLASS.ANIMAL two LOC yard  
 ‘There are two dogs in the yard.’ (Keo: Baird 2002; p140)
- (27) *ʔata ŋgaʔe dima ka dəra ndia*  
 person CLASS.HUM five eat day here  
 ‘Five people ate lunch here.’ (Keo: Baird 2002; p187)
- (28) *nio puʔu dima rua*  
 coconut CLASS.TREE five two  
 ‘seven coconut trees’ (Keo: Baird 2002; p146)
- (29) *sapi kami eko lima əsa*  
 cow 1PL.EXCL CLASS.ANIMAL five one  
 ‘We have six cows. (Lit: Our cows are six.)’ (Rongga: Arka et al. 2011; p36)
- (30) *mbo ito əsa rua ndau*  
 house small CLASS.GEN two DEM  
 ‘those two small houses’ (Rongga: Arka et al. 2011; p.xvii)
- (31) *kode fai mori rua*  
 person female CLASS.HUM two

‘those two women’ (Rongga: Arka et al. 2011; p7)

- (32) *dʒaʔo maŋa kamba eko wutu*  
 1SG have buffalo CLASS.ANIMAL four  
 ‘I have four buffaloes.’ (Rongga: Arka et al. 2011; p95)

All three of these languages have obligatory numeral classifiers, so I will mark all three of them with ‘+’ on the feature chart.

	Lio	Keo	Rongga
Numeral classifiers:	+	+	+

## 5.7 Verby adjectives

Mekong-Mamberamo languages tend to have ‘verby adjectives’: adjectives show similar morphosyntactic behaviour to verbs. This stands in contrast to languages where adjectives take either nominal marking, or their own special marking.

The Central Flores languages have verby adjectives. Due to Central Flores languages’ paucity of bound morphology, determining word classes is tricky and relies on distributional criteria. Nouns, adjectives and verbs can all serve as the head of a predicate in these languages. Baird (2002; p132) does not posit an adjective class separate from verbs for Keo, although Arka does for Rongga (2016; p118).

In the Central Flores languages, adjectives can be defined as a sub-class of verbs whose distinguishing characteristic is that they can modify an NP attributively without a relativiser. So, the Lio phrase *ae (eo) kəta ina* ‘that hot water’ in Lio can be expressed with or without the relativiser *eo*, since *aəta* ‘to be cold’ is in the sub-class of adjectival verbs:

- (33) *ae (eo) kəta ina*  
 water (REL) cold DEM  
 ‘that cold water’ (Lio: Elias 2017 fieldnotes)

However, in order to express ‘that running person’, where the noun ‘person’ is attributively modified by the non-adjectival verb ‘run’, the relativiser *eo* is required. Omitting it leads to ungrammaticality:

- (34) *ata eo paru ina*  
 person REL run DEM  
 ‘that running person’ (Lio: Elias 2017 fieldnotes)

- (35) \**ata paru ina*  
 person run DEM  
 Failed reading: ‘\*that running person’ (Lio: Elias 2017 fieldnotes)

Similarly Rongga *ata* is optional when used with an adjectival verb, but not with other non-adjectival verbs:

- (36) *naiŋge (ata) teʔa*  
 tamarind (REL) ripe  
 ‘ripe tamarind’ (Rongga: Arka et al. 2011; p52)

- (37) *ana ata mai ndau*  
 child REL come DEM  
 ‘the child that is coming’ (Rongga: Arka 2016; p284)

However, in Keo, Baird (2002; p409) mentions that relative clauses can omit the relativiser *ta* when the relative clause only contains one or two elements, but does not give any examples of a relative clause consisting of a non-adjectival verb with no relativiser.

Many concepts encoded as adjectives in English are expressed in Central Flores languages as verbs that can be used attributively, intransitively or transitively. Baird (2002; p132-133) uses the example of Keo *pətu* ‘hot, to heat’ to illustrate this fact. Analysing adjectives as a separate class from verbs would greatly complicate the analysis of the numerous cases such as these. The first sentence shows an intransitive predicative use, the second shows a transitive predicative use, and the third shows an attributive use.

- (38) *minu te pətu reʔe-reʔe*  
 drink DEM hot very:REDUP  
 ‘This drink is very hot.’ (Keo: Baird 2002; p132)

- (39) *rəke ha=goʔo ŋaʔo pətu ʔae*  
 wait SG=little 1SG hot water  
 ‘Wait a moment while I heat the water.’ (Keo: Baird 2002; p133)

- (40) *ŋaʔo minu kopi pətu*  
 1SG drink coffee hot  
 ‘I’m drinking hot coffee.’ (Keo: Baird 2002; p133)



All three of these languages have verby adjectives, so I will mark all three of them with ‘+’ on the feature chart.

	Lio	Keo	Rongga
Verby adjectives:	+	+	+

## 5.8 Basic SVO constituent order

One of the features typical of Mekong-Mamberamo languages is basic SVO constituent order in transitive clauses. Although SVO constituent order is cross-linguistically extremely common, all the neighbouring languages areas (South Asian, Northeast Asian, Taiwan/Philippines, New Guinea, Australia) have other dominant constituent orders.

In the Central Flores languages, the unmarked constituent order is indeed SVO. The following examples illustrate basic SVO sentences with full NP arguments.

- (41) *fua toki lako na*  
 wasp bite dog DEM  
 ‘The wasps are biting the dog. (Lio: Ibu\_Ferdy\_Frogstory)’

- (42) *Ardi pongga ana ndau*  
 Ardi hit child DEM  
 ‘Ardi hit the child.’ (Rongga: Arka et al. 2011; p.xv)

- (43) *?ana ke ngae kadzu*  
 child DEM search wood  
 ‘That child searched for wood.’ (Keo: Baird 2002; p82)

However, SVO clauses with two NP arguments are rare. Ellipsis of core arguments is extremely common in the Central Flores languages if the referent is clear from context. The following example shows ellipsis of the subject of an intransitive verb in Keo:

- (44) *bapa ena ?emba*  
 dad LOC where  
 ‘Where’s dad?’

*ø rio*  
*ø bathe*

‘(He’s) bathing.’ (Keo: Baird 2002; p274)

It is equally possible to ellipse the object of a transitive verb:

- (45) *na ʔemba sura ko ŋaʔo*  
LOC where letter GEN 1SG  
‘Where is my letter?’

*ine ŋatu ø*  
mother send ø

‘Mother sent (it).’ (Keo: Baird 2002; p275)

The following is an example from my Lio corpus of a transitive verb with both core arguments ellipsed:

- (46) *fua paru ŋai lako polu*  
wasp run because dog bark  
‘The wasps run because the dog is barking.’

*ø iwa toki ø*  
ø NEG bite ø

‘(The wasps) do not bite (the dog).’ (Lio: Ibu\_Ferdy\_Frogstory)

All three of these languages have basic SVO constituent order, so I will mark all three of them with ‘+’ on the feature chart.

	Lio	Keo	Rongga
Basic SVO constituent order:	+	+	+

## 5.9 Iamitive perfects

The iamitive aspect refers to the colexicalisation of two distinct but related concepts into a single aspectual category: 1) transitions into new states which still hold at the time of reference (the perfect) and 2) events which are completed and are viewed as a finished whole (the perfective). In English, the former might be expressed with an adjective plus

*already*, as in ‘I’m already full’, while the latter might be expressed with the past perfect, as in ‘I have eaten (already).’ On the other hand, in Malay, the equivalent sentences *Saya sudah kenyang* ‘I’m already full’ and *Saya sudah makan* ‘I have (already) eaten’ are formally identical.

In Lio and Rongga too, these two senses are expressed in the same way, and hence the iamitive aspect is present as an aspectual category in Lio and Rongga.

- (47) *aku boʔo dowa*  
 1SG full PERF  
 ‘I am already full.’ (Lio: Elias 2017 fieldnotes)

- (48) *aku ka dowa*  
 1SG eat PERF  
 ‘I have eaten.’ (Lio: Elias 2017 fieldnotes)

- (49) *somo mbuʔe ga sia hongga ga*  
 because adult PERF 3PL handsome PERF  
 ‘Because they are already grown-up and handsome.’ (Rongga: Arka 2016; p278)

- (50) *kau dadi ga ana ndau*  
 2SG give.birth PERF child DEM  
 ‘You have given birth to that child.’ (Rongga: Arka 2016; p110)

However, Baird (2002) reports that Keo does distinguish between these two aspectual categories through the placement of the morpheme *neya*, *ya*. When placed before the predicate, it yields the ‘persistent perfect’ (a transition to a state that still holds at the time of speaking) but when placed after the predicate, it yields the ‘perfective/completive’ aspect (event viewed as a finished whole). To illustrate the difference, Baird (2002; p308) provides the following pair of examples:

- (51) *aʔi ʔaʔo neya poʔi. ʔaʔo mbana tado*  
 leg 1SG PER.PER break.1SG walk cannot  
 ‘My leg is broken. I can’t walk.’ (Baird: Baird 2002; p308)

- (52) *aʔi ʔaʔo poʔi neya. ʔaʔo bia poʔi wadi*  
 leg 1SG break PER.COM. 1SG not.want break again  
 ‘My leg has been broken. I don’t want it broken again!’ (Baird: Baird 2002; p308)

Thus, while the two senses of the iamitive are encoded by a single morpheme in Keo *neya*, *ya*, there is still a formal distinction between the two senses. Therefore, I assign Keo a ‘+/-’ to indicate that it displays some of the features of iamitive perfects.

	Lio	Keo	Rongga
Iamitive perfects:	+	+/-	+

### 5.10 ‘Give’ causatives

In Mekong-Mamberamo languages, causative constructions are often expressed with a morpheme identical to or derived from the verb ‘to give’. This is found in eastern Malay varieties such as Papuan Malay, where causatives like Standard Indonesian *mematikan* ‘to kill’ are often expressed as *kasi mati* ‘give die’ instead.

In Lio, the most common causative serial verb construction indeed uses the verb *pati* ‘to give’.

- (53) *guru pati duke kami ləka nia ana kelas satu*  
 teacher give kneel 1PL.INC LOC face child class one  
 ‘Teacher made us kneel in front of the first grade children.’ (Lio: Ibu\_Ferdy\_Scorpion\_Story)

Less commonly, the verb *tau* ‘to make’ also serves as a causativiser in Lio.

- (54) *ana mo tau masa nia*  
 child PROS make clean face  
 ‘The child is going to wash its face.’ (Lio: Ibu\_Ferdy\_Bridewealth)

However, in Keo and Rongga, only the verb *tau* ‘to make’ is used in serial verb constructions to express causation.

- (55) *selu tau mata manu ndau*  
 Selus make die chicken DEM  
 ‘Selus kills that chicken.’ (Rongga: Arka 2016; p227)

- (56) *?imu tau buge ?ana ?imu*  
 3SG make fat child 3SG  
 ‘She fattened her child.’ (Keo: Baird 2002; p118)

Keo and Rongga are both lacking ‘give’ causatives, so I will mark them with ‘-’ on the feature chart.

	Lio	Keo	Rongga
‘Give’ causatives:	+	-	-

### 5.11 Low differentiation of adnominal attributive constructions

This feature refers to the formal similarity of three types of adnominal attributive constructions: genitival, adjectival and relative clause constructions. In English, these three types of noun phrases are distinct on the surface: genitival (‘Adam’s book’), adjectival (‘the red book’) and relative (‘the book that Adam bought’). In the Mekong-Mamberamo area, languages tend to collapse these three syntactic constructions to some degree. Thus, in Minangkabau, the possessive, genitival and relative relations can all be expressed with simple juxtaposition:

- (57) *rumah fadʒar*  
house Fajar  
‘Fajar’s house’ (Minangkabau, Western Indonesia: Gil 2015; p292)

- (58) *rumah ketek*  
house small  
‘small house’ (Minangkabau, Western Indonesia: Gil 2015; p292)

- (59) *rumah fadʒar bali*  
house Fajar buy  
‘the house that Fajar bought’ (Minangkabau: Gil 2015; p292)

As seen in section 5.7, adjectival notions are expressed in the Central Flores languages by means of adjectival verbs, which may modify nouns with or without an intervening relativiser:

- (60) *ae (eo) kəta ina*  
water (REL) cold DEM  
‘that cold water’ (Lio: Elias 2017 fieldnotes)

- (61) *naŋge (ata) teʔa*  
tamarind (REL) ripe  
‘ripe tamarind’ (Rongga: Arka et al. 2011; p52)

- (62) *baki nio (ta) wadzo*  
 plank coconut (REL) old  
 ‘old coconut planks’ (Keo: Baird 2002; p410)

As for relative clauses with non-adjectival verbs Lio and Rongga, the relativiser (Lio *eo*, Rongga *ata*) is not optional:

- (63) *lako eo kai gəti*  
 dog REL 3SG buy  
 ‘the dog that he bought’ (Lio: Elias 2017 fieldnotes)

- (64) *ana ata mai ndau*  
 child REL come DEM  
 ‘that child who is coming’ (Rongga: Arka 2016; p284)

On the other hand, the relativiser *ta* is optional in relative phrases in Keo, and the following example is equally grammatical with and without it:

- (65) *puʔu kadzu (ta) ŋara mona nde*  
 trunk wood (REL) name NEG DEM  
 ‘that tree with no name’ (Keo: Baird 2002; p410)

In Lio, Keo and Rongga, adnominal attributive possession can be expressed by simple possessum-possessor juxtaposition:

- (66) *kolo lako*  
 head dog  
 ‘the dog’s head’ (Lio: Ibu.Ferdy.Frogstory)
- (67) *aʔi medza*  
 leg table  
 ‘the leg of the table’ (Keo: Baird 2002; p214)
- (68) *uma simeon*  
 garden Simeon  
 ‘Simeon’s garden’ (Rongga: Arka 2016; p188)

In Lio (but not in Keo and Rongga) it is possible to use the relativiser *eo* to express adnominal possession:

- (69) *taŋgo eo ata fai*  
 portion REL person female  
 ‘the woman’s portion’ (Lio: Ibu\_Ferdy\_Bridewealth)

In Rongga and Keo (but not in Lio) there is a genitive morpheme that sometimes appears between the possessum and the possessor: Keo *koʔo*, Rongga *ko*. In Keo, this particle is obligatory in some contexts, such as when the two nouns in the possessive adnominal construction are common nouns or kin terms. The genitive particle is also preferred in some contexts in Rongga, but not the same set of contexts as in Keo.

- (70) *ʔana koʔo wəta*  
 child GEN sister  
 ‘sister’s child’ (Keo: Baird 2002; p214)

- (71) *lako ko domi*  
 dog GEN Domi  
 ‘Domi’s dog’ (Rongga: Arka 2016; p187)

In summary, all three languages can use simple juxtaposition to express adnominal possession. In addition, Lio can use the relativiser *eo*, and Keo and Rongga can use the genitive marker Keo *koʔo*, Rongga *ko* to express adnominal possession.

Overall, the picture is mixed. Lio has the lowest level of differentiation of adnominal constructions, because the relativiser *eo* can be used to form possessive, genitive and relative constructions. Keo also has a low level of differentiation of adnominal constructions, since juxtaposition is used to express possessive, genitive and relative constructions. However, in some classes of genitive constructions, the genitive particle Keo *koʔo* is obligatory, so some types of possession cannot be expressed through juxtaposition. Therefore, I assign Keo a ‘+/-’ in the table of features. Rongga has a higher level of differentiation of adnominal constructions, since it has not only the genitive particle *ko* in many genitive constructions, it also cannot generally drop the relativiser *ata* in relative clauses. Therefore I assign Rongga a ‘-’ in the feature chart.

	Lio	Keo	Rongga
Low differentiation of adnominal attributive constructions:	+	+/-	-

## 5.12 Weakly developed grammatical voice

A language can be said to have a weakly developed grammatical voice system if there is no overt, morphologically marked mechanism for voice alternations such as the passive. This is generally true of the Mekong-Mamberamo languages. Voice alternations encoded only as constituent order changes still qualify as ‘weakly developed grammatical voice’ by the definition of Gil (2015).

The Central Flores languages do not have dedicated morphology for voice alternations and therefore show weakly developed grammatical voice. A system of voice alternation through constituent order changes is grammaticalised to the greatest degree in Rongga, where passive is systematically expressed by promoting the Patient to subject position, and reintroducing the Agent in a prepositional phrase with *ne* ‘with, by’.

- (72) *ardi ponga ana ndau*  
 Ardi hit child DEM  
 ‘Ardi hit that child. (\*That child hit Ardi.)’ (Rongga: Arka 2016; p217)

- (73) *ana ndau ponga ne ardi*  
 child DEM hit by Ardi  
 ‘That child was hit by Ardi.’ (Rongga: Arka 2016; p217)

Rongga also shows the following restriction on relative clauses: the object of an active transitive clause cannot be relativised, but must be reformulated as the subject of a passive clause first.

- (74) *\*ana ata ardi ponga ndau bako dʒaʔo*  
 child REL Ardi hit DEM nephew 1SG  
 Failed reading: ‘\*The child that Ardi hit is my nephew.’ (Rongga: Arka 2016; p220)

- (75) *ana ata ponga ne ardi ndau bako dʒaʔo*  
 child REL hit by Ardi DEM nephew 1SG  
 Acceptable: ‘The child that was hit by Ardi is my nephew.’ (Rongga: Arka 2016; p220)

Lio and Keo are lacking the grammaticalised voice alternation system described for Rongga. In Keo, the object in an active SVO sentence can be fronted to topicalise it, but the Patient remains the object, and the Agent remains the subject. The Agent is not demoted to become an oblique argument, and therefore this is not a true passive construction. Lio also lacks a true passive, but uses the same object fronting strategy as Keo. The



first example shows a regular active SVO sentence where ‘Nus’ is the subject/Agent and ‘Arno’ is the object/Patient:

- (76) *nus boʻa arno*  
 Nus hit Arno  
 S V O  
 ‘Nus hit Arno.’ (Keo: Baird 2002; p78, modified)

When the object fronted, this yields the following OSV clause, but ‘Arno’ remains the object/Patient despite it being in initial position. The subject/Agent ‘Nus’ is still required and has not been demoted.

- (77) *arno nus boʻa*  
 Arno Nus hit  
 O S V  
 ‘Nus hit Arno. (Not: \*Arno was hit by Nus.)’ (Keo: Baird 2002; p78)

The final example shows that it is not a true passive construction because the following sentence must be interpreted as an active SV(O) clause with the object elided, rather than a passive SV clause with the Patient as subject.

- (78) *arno boʻa*  
 Arno hit  
 S V  
 ‘Arno hit (someone). (Not: \*Arno was hit.)’ (Keo: Baird 2002; p79)

Unlike in Rongga, there are no restrictions on relativising the object of an active clause in Lio and Keo:

- (79) *dəlu aku manusia eo fua toki*  
 friend 1SG person REL wasp bite  
 ‘My friend is the person who the wasps bit.’ (Lio: Ibu\_Ferdy\_Frogstory)

- (80) *?ata ta ?imu boʻa ke palu*  
 person REL 3SG hit DEM run  
 ‘The person that he hit ran.’ (Keo: Baird 2002; p72)

All three of these languages have weakly developed grammatical voice, so I will mark all three of them with ‘+’ on the feature chart.

	Lio	Keo	Rongga
Weakly developed grammatical voice:	+	+	+

### 5.13 Isolating word structure

Mekong-Mamberamo languages typically have a low number of morphemes per word, with grammatical morphemes expressed as independent words.

This is very much true of the Central Flores languages. They are notable precisely for their almost total absence of bound affixes, in contrast to the modest amount of morphology present in the languages of West and East Flores (see Sections 4.1, 4.2).

There is one proclitic (Lio and Rongga *sa=*, Keo *ha=*) which marks the singular number in numeral phrases and is reconstructible as Proto-Central Flores *\*sa=* (Elias 2018). This is transparently related to the form of the numeral for one, PCF *\*əsa*.

All three of these languages have isolating word structure, so I will mark all three of them with ‘+’ on the feature chart.

	Lio	Keo	Rongga
Isolating word structure:	+	+	+

### 5.14 Short words

Mekong-Mamberamo languages typically have short words, which is unsurprising given their isolating morphology. In many languages, there are constraints on the maximum size of a word.

The Central Flores languages have a very restricted range of possible word shapes. They allow only open syllables and do not allow any consonant clusters, with maximally disyllabic word of form CVCV. Native, monomorphemic words do exceed two syllables (see Table 9 for an exhaustive list of possible word shapes in Central Flores). Words longer than two syllables are either loans, or formed by compounding or fossilisation of the singular proclitic (Lio, Rongga *sa=*, Keo *ha=*).

Shape	Lio	Keo	Rongga
V	e ‘think’	e ‘think’	e ‘exclamation’
CV	ka ‘eat’	fu ‘hair’	ʃa ‘plate’
VV	ae ‘water’	oa ‘request’	ua ‘rattan’
VCV	eko ‘tail’	uwa ‘skin’	aŋe ‘maybe’
CVV	ria ‘large’	loa ‘burn’	lea ‘ginger’
CVCV	pati ‘give’	rəto ‘dip’	təlo ‘egg’

Table 9: Exhaustive list of possible word shapes in Central Flores languages

All three of these languages have short words, so I will mark all three of them with ‘+’ on the feature chart.

	Lio	Keo	Rongga
Short words:	+	+	+

### 5.15 Low grammatical-morpheme density

Mekong-Mamberamo languages often display a low grammatical-morpheme density. Utterances are often composed mainly of lexical items, with few other morphemes required to bind them into a grammatical utterance. This is logically distinct from isolating morphology, but often co-occurs with it.

The Central Flores languages have low grammatical morpheme density. Since very few semantic distinctions are obligatorily expressed on either verbs or nouns, rather long sentences consisting only of content words are not uncommon in the Central Flores languages.

- (81) *ata fai kodo tei ana wawi*  
 person female look find child pig  
 ‘The woman looks and sees a piglet. (Lio: Elias 2017 fieldnotes)’

This sentence requires the use of 5 grammatical morphemes in English: each noun phrase must receive either a definite article or an indefinite article, each 3SG verb must take a final /-s/, and conjunction must be expressed using *and*. None of these are required in Lio, which expresses the notion of ‘looking and seeing’ as a bare serial verb construction and does not obligatorily express definiteness on nouns or tense on verbs. The same holds true in the other Central Flores languages.

All three of these languages have low grammatical-morpheme density, so I will mark all three of them with ‘+’ on the feature chart.

	Lio	Keo	Rongga
Low grammatical morpheme density:	+	+	+

### 5.16 Optional thematic-role flagging

A feature of the Mekong-Mamberamo language area is optional thematic-role flagging, which means that the arguments of a verb are not necessarily overtly marked to indicate their relationship to the verb. Cross-linguistically, oblique arguments are more likely to require an overt marker (such as a preposition) than core arguments.

Gil (2015) cites examples from languages in the Mekong-Mamberamo area where oblique arguments do not need to be introduced by an overt marker (‘bare oblique’ constructions), such as the following:

- (82) *dʒələma dahar taŋkal*  
 person eat tree  
 ‘The man is eating by the tree.’ (Sundanese, Western Indonesia: Gil 2015; p317)

- (83) *isok et mega*  
 man eat tree  
 ‘The man is eating by the tree.’ (Meyah, Western New Guinea: Gil 2015; p317)

While Central Flores languages do not overtly mark the core arguments of the verb, marking of oblique arguments is obligatory. Thus, the following Lio sentence is ungrammatical without the use of *ləka* ‘in, at’ to introduce the oblique argument:

- (84) *ani mərə ləka puʔu kadʒu*  
 bee live LOC trunk wood  
 ‘The bees live in the tree.’ (Lio: Ibu\_Ferdy\_Frogstory)

- (85) *\*ani mərə puʔu kadʒu*  
 bee live trunk wood  
 Failed reading: ‘\*The bees live in the tree.’ (Lio: Ibu\_Ferdy\_Frogstory)

Oblique arguments must be overtly marked in Keo and Rongga as well.

- (86) *ʔimu kere dau maʔu*  
 3SG wait down beach  
 ‘He waited down at the beach.’ (Keo: Baird 2002; p89)

- (87) *kari ŋgoe one radʔi*  
 3SG fall LOC stair  
 ‘He fell on the stairs.’ (Rongga: Arka 2016; p144)

Since the Central Flores languages lack the bare oblique construction Gil (2015) cites in other Mekong-Mamberamo languages, I give all three languages a ‘-’ for this feature.

	Lio	Keo	Rongga
Optional thematic-role flagging:	-	-	-

### 5.17 Optional TAM marking

In many Mekong-Mamberamo languages, the expression of tense, aspect and mood is optional.

In the Central Flores languages, there is not obligatory marking for tense, aspect and mood. A clause with no overt TAM marking is not restricted in its range of possible interpretations. Hence, the following Lio sentence with no TAM marking could receive a range of possible temporal interpretations:

- (88) *fua toki lako na*  
wasp bite dog DEM  
‘The wasps [are biting/bit] the dog.’ (Lio: Ibu\_Ferdy\_Frogstory)

TAM marking is also not obligatory in Keo or Rongga:

- (89) *kaʔe ka ʔuwi dʒawa*  
older.sibling eat tuber Java  
‘Big brother [is eating/ate] sweet potato.’ (Keo: Baird 2002; p78)

- (90) *pondo ndau mado*  
pot DEM fall  
‘That pot [is falling/fell].’ (Rongga: Arka 2016; p157)

All three of these languages have optional TAM marking, so I will mark all three of them with ‘+’ on the feature chart.

	Lio	Keo	Rongga
Optional TAM marking:	+	+	+

### 5.18 Summary: Mekong-Mamberamo features in Central Flores languages

I have examined the 17 features identified by Gil (2015) as typical of the Mekong-Mamberamo linguistic area and assessed their presence or absence in the languages of Central Flores. The following table summarises the findings:

	Lio	Keo	Rongga
The passing gesture:	+	?	?
Repeated dental clicks expressing amazement:	+	?	?
Conventionalised greeting with ‘where’:	+	+	?
‘eye day’ to ‘sun’ lexicalisation:	+	+	+
d/t place-of-articulation mismatch:	+	?	?
Numeral classifiers:	+	+	+
Verby adjectives:	+	+	+
Basic SVO constituent order:	+	+	+
Iamitive perfects:	+	+/-	+
‘Give’ causatives:	+	-	-
Low differentiation of adnominal attributive constructions:	+	+/-	-
Weakly developed grammatical voice:	+	+	+
Isolating word structure:	+	+	+
Short words:	+	+	+
Low grammatical morpheme density:	+	+	+
Optional thematic-role flagging:	-	-	-
Optional TAM marking:	+	+	+

The only feature which is missing from all three languages is ‘optional thematic-role flagging’. In all three languages in my sample, oblique arguments are marked obligatorily, not optionally.

## 6 Additional evidence from the Central Flores numeral system

In this section, I describe the multiple numeral systems found in Lio, Keo and Rongga, which shed light on certain aspects of the pre-Austronesian languages of Central Flores.

One of the defining innovations of the Central Flores languages is the restructuring of the Austronesian decimal numeral system into a mixed-base quinary-decimal system. The following table (Table 10) shows the decimal PMP numerals on the left, followed by languages of East Flores (Sika) and West Flores (Manggarai) which retained that system more or less intact. Also included are the Palu’e numerals, to show that these are of the decimal Austronesian type and not the quinary-decimal Central Flores type. These are contrasted with the reconstructed Proto-Central Flores numerals (Elias 2018) and their reflexes in Lio, Keo and Rongga in Table 12.

The Proto-Central Flores numerals (plus the obligatory general classifier \*əsa) are reconstructed in Elias (2018) as follows: \*sa=[əsa] ‘one’, \*[əsa] dua two, \*[əsa] təlu three, \*[əsa] wutu four, \*[əsa] lima five, \*[əsa] lima əsa six, \*[əsa] lima dua seven, \*[əsa] dua mbutu

eight, \*[əsa] tʔra əsa nine. The higher bases are \*mbulu ten, \*ŋasu one hundred, \*riwu one thousand.

The numerals from 6 to 9 are derived from the lower numerals by a number of different strategies. PCF \*lima əsa ‘six’ means ‘one five’ and PCF \*lima dua ‘seven’ means ‘five two’, using an additive strategy. PCF \*dua mbutu ‘eight’ means ‘two four’ using a multiplicative strategy (note the prenasalisation on the second element compared with PCF \*wutu ‘four’; this may be the remnant of the PMP \*-ŋa- morpheme which appears in PMP \*sa-ŋa-puluq, PCF \*sa=mbulu ‘ten’). Finally, PCF \*təra əsa ‘nine’ seems to be composed of an initial morpheme meaning ‘to take away, to remove’ followed by ‘one’, so it means something like ‘take away one (from ten)’ using a subtractive strategy.

Note that the general classifier PCF \*əsa is homophonous with the morpheme used to represent the number 1 in the composed numerals 6 and 9 (and may share an etymology in PMP \*əsa ‘one’, although they are now clearly separate in Central Flores). In order to help the reader, the instances where PCF \*əsa is used as a classifier are placed between square brackets. If the Central Flores speaker were counting something that requires a different classifier, such as animals, all instances of PCF \*əsa ‘general classifier’ would be replaced with PCF \*eko ‘animal classifier’ (Lio, Rongga *eko*, Keo *?eko*).

	PMP	Sika	Manggarai	Palu'e
1	*esa	ha	ʈa	a
2	*duha	rua	sua	rua
3	*telu	təlu	təlu	təlu
4	*epat	hutu	pat	ʃa
5	*lima	lima	lima	lima
6	*enem	əna	ənəm	ʔəne
7	*pitu	pitu	pitu	ʃitu
8	*walu	walu	alo	walu
9	*siwa	hiwa	ʈiok	iwa
10	*sa-ŋa-puluq	pulu ha	ʈa mpulu	a pulu
11	-	pulu wot ha	ʈa mpulu ʈa	-
12	-	pulu wot rua	ʈa mpulu sua	-
13	-	pulu wot təlu	ʈa mpulu təlu	-
14	-	pulu wot hutu	ʈa mpulu pat	-
15	-	pulu wot lima	ʈa mpulu lima	-
16	-	pulu wot əna	ʈa mpulu ənəm	-
17	-	pulu wot pitu	ʈa mpulu pitu	-
18	-	pulu wot walu	ʈa mpulu alo	-
19	-	pulu wot hiwa	ʈa mpulu ʈiok	-
20	*duha-ŋa-puluq	pulu rua	sua mpulu	rua pulu
100	*sa-ŋa-Ratus	ŋasu ha	ʈa ratus	a ʈatu
1000	*sa-ŋa-Ribu	riwu ha	sa=səbu	a riwu

Table 10: Inherited decimal numerals in the languages of East Flores, West Flores and Palu'e



	PCF
1	*sa=[əsa]
2	*[əsa] dua
3	*[əsa] təlu
4	*[əsa] wutu
5	*[əsa] lima
6	*[əsa] lima əsa
7	*[əsa] lima dua
8	*[əsa] dua mbutu
9	*[əsa] təra əsa
10	*sa=mbulu
11	*sa=mbulu sa=[əsa]
12	*sa=mbulu [əsa] dua
13	*sa=mbulu [əsa] təlu
14	*sa=mbulu [əsa] wutu
15	*sa=mbulu [əsa] lima
16	*sa=mbulu [əsa] lima əsa
17	*sa=mbulu [əsa] lima dua
18	*sa=mbulu [əsa] dua mbutu
19	*sa=mbulu [əsa] təra əsa
20	*mbulu dua
100	*sa=ŋasu
1000	*sa=riwu

Table 11: Quinary-decimal numerals in Proto-Central Flores

	Lio	Keo	Rongga
1	sa=[əsa]	ha=?əsa	sa=[əsa]
2	[əsa] rua	?əsa rua	[əsa] .lua
3	[əsa] təlu	[?əsa] tədu	[əsa] təlu
4	[əsa] sutu	[?əsa] wutu	[əsa] wutu
5	[əsa] lima	[?əsa] dima	[əsa] lima
6	[əsa] lima əsa	[?əsa] dima əsa	[əsa] lima əsa
7	[əsa] lima rua	[?əsa] dima rua	[əsa] lima .lua
8	[əsa] rua mbutu	[?əsa] rua mbutu	[əsa] .lua mbutu
9	[əsa] təra əsa	[?əsa] təra ?əsa	[əsa] təra əsa
10	sa=mbulu	ha=mbudu	sa=mbulu
11	sa=mbulu sa=[əsa]	ha=mbudu ha=[?əsa]	sa=mbulu sa=[əsa]
12	sa=mbulu [əsa] rua	ha=mbudu [?əsa] rua	sa=mbulu [əsa] .lua
13	sa=mbulu [əsa] təlu	ha=mbudu [?əsa] tədu	sa=mbulu [əsa] təlu
14	sa=mbulu [əsa] sutu	ha=mbudu [?əsa] wutu	sa=mbulu [əsa] wutu
15	sa=mbulu [əsa] lima	ha=mbudu [?əsa] dima	sa=mbulu [əsa] lima
16	sa=mbulu [əsa] lima əsa	ha=mbudu [?əsa] dima ?əsa	sa=mbulu [əsa] lima əsa
17	sa=mbulu [əsa] lima dua	ha=mbudu [?əsa] dima rua	sa=mbulu [əsa] lima .lua
18	sa=mbulu [əsa] rua mbutu	ha=mbudu [?əsa] rua mbutu	sa=mbulu [əsa] .lua mbutu
19	sa=mbulu [əsa] təra əsa	ha=mbudu [?əsa] təra ?əsa	sa=mbulu [əsa] təra əsa
20	mbulu rua	mbudu rua	mbulu .lua
100	sa=ŋasu	ha=ŋasu	sa=ŋasu
1000	sa=riwu	ha=liwu	sa=riwu

Table 12: Quinary-decimal numerals in the modern languages of Central Flores

In addition to the quinary-decimal system, there is a quaternary (base 4) numeral system which is present in all Central Flores languages and is therefore reconstructible to Proto-Central Flores. This system has a more restricted application, typically being used (by Lio speakers at least) when dealing with small objects which can be stacked into pyramids of 4 such as coconuts, areca nuts, or limes.

Interestingly, the highest repeating base in this system (40) is reconstructed as PCF \*ulu ‘head; 40’. Non-Austronesian (Papuan) languages of the region are well-known for using body part words as numerical bases (Schapper and Klamer 2014), so this may reflect semantic influence from a non-Austronesian language.

	PCF	Lio	Keo	Rongga
1	*sa=[əsa]	sa=[əsa]	ha=[ʔəsa]	sa=[əsa]
2	*[əsa] dua	[əsa] rua	[ʔəsa] rua	[əsa] .rua
3	*[əsa] təlu	[əsa] təlu	[ʔəsa] tədu	[əsa] təlu
4	*sa=liwu	sa=liwu	ha=diwu	sa=liwu
5	*sa=liwu sa=[əsa]	sa=liwu sa=[əsa]	ha=diwu ha=[ʔəsa]	sa=liwu sa=[əsa]
6	*sa=liwu [əsa] dua	sa=liwu [əsa] rua	ha=diwu [ʔəsa] rua	sa=liwu [əsa] .rua
7	*sa=liwu [əsa] təlu	sa=liwu [əsa] təlu	ha=diwu [ʔəsa] tədu	sa=liwu [əsa] təlu
8	*liwu dua	liwu rua	diwu rua	liwu .rua
9	*liwu dua sa=[əsa]	liwu rua sa=[əsa]	diwu rua ha=[ʔəsa]	liwu .rua sa=[əsa]
10	*liwu dua [əsa] dua	liwu rua [əsa] rua	diwu rua [ʔəsa] rua	liwu .rua [əsa] .rua
...	...	...	...	...
40	*sa=ulu	sa=ulu	ha=ʔudu	sa=ulu

Table 13: Quaternary numerals in the languages of Central Flores

Finally, Arka (2016; p127-128) describes an intriguing additional decimal numeral system in Rongga, which brings the number of distinct numeral systems in that language to three, with three separate bases (quaternary, quinary, decimal). This system is not productive in that it does not go above 20 and the formation of numbers above 10 is not transparent. A notable feature of this counting system is that various numbers contain consonant clusters which violate the phonotactic rules of Rongga: particularly *ŋgwo* ‘nine’ and *mopla* ‘twenty’. Furthermore, none of the numerals below ten except *dwa* ‘two’ and possibly *da* ‘one’ have a plausible Austronesian etymology. The numerals 3-10 do not resemble the numerals in any Austronesian language, nor in any of the nearby non-Austronesian Timor-Alor-Pantar languages (Schapper and Klamer 2014). The formation of the numerals 11-15 is intriguing: here we see the inherited Central Flores etyma PCF \*əsa ‘one’, \*dua ‘two’, \*təlu ‘three’, \*wutu ‘four’, \*lima ‘five’ make an appearance. It seems that the strategy for forming numerals 11-15 is by reduplicating the numerals for 1-5 with an unpredictably altered initial consonant. Thus, while the numerals *dəmu* ‘three’, *dəke* ‘four’, *ali* ‘five’ do not reflect PCF numerals, the numerals *təlungətu* ‘thirteen’, *wutunggutu* ‘fourteen’ and *limakima* ‘fifteen’ do reflect the PCF numerals. From 16 to 20, almost nothing can be ascertained about the etymology of the numerals except that *watopəsa* ‘nineteen’ may perhaps contain a morpheme reflecting PCF əsa ‘one’ and therefore be formed by subtraction (20 - 1) (Arka 2016; p127).

Throughout this number system, there is a pervasive tendency for neighbouring numerals to alliterate with each other. This is a commonly noted development cross-linguistically. In the Rongga numeral system under discussion, there seems to be frequent use of such alliteration: note the pairs *sipi* ‘seven’, *sapa* ‘eight’, or the pair *ŋgwo* ‘nine’, *ŋguru* ‘ten’. The numeral *təlungətu* ‘thirteen’ appears to have shifted from the expected \*\**təlungəlu*

under the influence of the following *wutungutu* ‘fourteen’, so that the two numerals share the same ending.

	Rongga
1	ɖa
2	ɖua
3	dəmu
4	ɖəke
5	ali
6	woe
7	sipi
8	sapa
9	ŋgwo
10	ŋguru
11	əsəŋgəsa
12	ɾuaɖua
13	təlunɡətu
14	wutunɡutu
15	limakima
16	aŋɡunae
17	nəŋɡonae
18	soroila
19	watopəsa
20	mopla

Table 14: Non-productive numerals in Rongga

This non-productive Rongga numeral system shares a number of suggestive parallels with the ‘Yan Tan Tethera’ sheep-counting systems of Northern England (Ingram 1977), which are a remarkable instance of lexical items from a substrate language enduring for many centuries.

The sheep-counting systems of Northern England were used by shepherds into the early 20th century. They are limited to numbers below 20 (like Rongga) and are derived from the extinct Brythonic Celtic languages of Northern England such as Cumbric which form a linguistic substrate in this part of England. The numerals somewhat resemble those found in modern Brythonic languages such as Welsh, but adjacent numerals are often altered to alliterate better (like in Rongga). In the regions of Northern England where these sheep-counting systems are used, the Celtic languages have been extinct for centuries: for instance, Cumbric itself was extinct by the 12th century. However, the sheep-counting systems survived into the 20th century (in modified form) as a last vestige of the earlier Celtic languages. Several Northern English variants of the sheep-counting system are

presented, along with modern Welsh for comparison (Ingram 1977).

In my view, the non-productive Rongga counting system laid out in Table 14 is analogous to the sheep-counting systems of Northern England: the last lexical remnant of a group of long-extinct substrate languages. Based on the lack of similarities between the forms of the numerals and any other languages of the area, I suggest that the substrate language has no living descendants or relatives in the area.

	Welsh	Bowland	Coniston	Tong
1	un	yain	yan	yan
2	dau	tain	taen	tan
3	tri	eddera	tedderte	tether
4	pedwar	peddera	medderte	methier
5	pump	pit	pimp	pick
6	chwech	tayter	haata	sesan
7	saith	layter	slaata	asel
8	wyth	overa	lowra	catel
9	naw	covera	dowra	oiner
10	deg	dix	dick	dick
11	un ar ddeg	yain-a-dix	yan-a-dick	yanadick
12	deuddeg	tain-a-dix	taen-a-dick	tanadick
13	tri ar ddeg	eddera-a-dix	tedder-a-dick	tetheradick
14	pedwar ar ddeg	peddera-a-dix	medder-a-dick	methieradick
15	pymtheg	bumfit	mimph	bumfit
16	un ar bymtheg	yain-a-bumfit	yan-a-mimph	yanabum
17	dau ar bymtheg	tain-a-bumfit	taen-a-mimph	tanabum
18	deunaw	eddera-bumfit	tedder-a-mimph	tetherabum
19	pedwar ar bymtheg	peddera-a-bumfit	medder-a-mimph	methierabum
20	ugain	jiggit	gigget	jigget

Table 15: Borrowed Celtic sheep-counting systems used in Northern England

## 7 Historical proposal for Central Flores languages

The historical proposal that follows is informed by Thomason and Kaufman (1988)’s framework of contact scenarios and their linguistic outcomes. They draw a distinction between two basic types of contact-induced change: ‘borrowing’ and ‘substratum interference’.

Borrowing refers to ‘the incorporation of foreign features into a group’s native language by speakers of that language: the native language is maintained but is changed by the addition of the incorporated features’ (Thomas and Kaufman 1988; p37). The first changes are lexical adoptions, and with more intense contact phonological, morphological and syntactic

elements may be borrowed as well. Lexical borrowing may take place without widespread societal bilingualism, while grammatical borrowing usually takes place in a situation of extensive bilingualism. An example Thomason and Kaufman (1988) give is the influence of Sanskrit on the Mainland Southeast Asian languages: while many Sanskrit words have been adopted into Thai, Khmer and other SE Asian languages, their grammatical structure has not been changed much and the bulk of Thai and Khmer speakers were never proficient in Sanskrit. Another example they give is the effect of Indian languages on the English of the British colonists. English adopted local words for concepts which did not exist before, but the grammar was hardly affected.

Substratum interference, unlike borrowing, occurs when ‘a group of speakers shifting to a target language fails to learn the target language correctly’ (Thomas and Kaufman 1988; p39). Unlike in borrowing, in substratum interference the group of speakers does not maintain their own native language, but adopts the native language of another group with which they are in close contact. The linguistic traces of substratum interference are very different from borrowing: ‘unlike borrowing, interference through imperfect learning does *not* begin with vocabulary: it begins instead with sounds and syntax, and sometimes includes morphology as well before words from the shifting group’s original language appear in the T[arget] L[anguage]’ (Thomason and Kaufman 1988; p39, emphasis in original). The grammatical effects of substratum interference are particularly strong when the shifting group is numerically greater than the target language speakers. Thus, the linguistic effects of substratum interference are almost opposite to those of borrowing: structural interference first, then vocabulary. Often, the target language is of higher prestige than the shifting language, and hence the vocabulary of the target language will be preferred over the speakers’ native vocabulary. An example of substratum interference given by Thomason and Kaufman (1988) is the case of Dravidian influence in Sanskrit. The Dravidian languages are held to be the source of many of the distinctive typological features of Sanskrit (retroflex consonants, among others) but there is not a great amount of clearly Dravidian vocabulary in Sanskrit. This is because Dravidian speakers shifted to Sanskrit, rather than Sanskrit speakers borrowing from Dravidian.

Another example of substratum interference is the effect of Uralic on the Slavic languages. Uralic features in Slavic are held to include phonemic palatalisation, large case systems and the use of the partitive genitive, among others. However, some have argued that since there is very little evidence of Uralic vocabulary in Slavic, this is evidence against close contact between them. In fact, this does not argue against contact: it merely argues against a situation where Slavic borrowed heavily from Uralic. It is completely consistent with a substratum interference scenario where Uralic speakers shifted to Slavic, resulting in grammatical but not lexical interference.

Another difference between borrowing and substratum interference is the amount of time required for far-reaching structural modification. For extensive grammatical borrowing to take place, usually a long period of time is required. However, substratum interference may take place in a single generation: ‘in fact, substratum features are more likely to enter

a T[arget] L[anguage] rapidly than slowly: if the shift takes place over long centuries, then the shifting population is likely to be truly bilingual in the T[arget] L[anguage]’ (Thomason and Kaufman 1988; p41). Hence, the strongest effects of substratum interference will be seen in cases where the transition happened most abruptly.

In the case of Central Flores, a substratum interference scenario is clearly preferred. The Central Flores languages show an almost total restructuring of their typological profile, but are lexically very conservative. Comparing a 100-item Swadesh vocabulary list of PMP with Lio, Keo and Rongga yielded the following result: 69% retentions from PMP in Lio, 64% retentions in Keo and 63% retentions in Rongga. A high level of lexical conservatism, combined with heavy grammatical interference, is diagnostic of a substratum interference scenario brought about by rather abrupt language shift according to Thomas and Kaufman (1988).

With this in mind, I propose the following historical scenario to account for the typological profile of the Central Flores languages: around 3000 years before the present, the Austronesians began their push into the Indonesian archipelago from the Philippines. They brought with them richly inflected VSO Austronesian languages similar to the modern languages of Taiwan and the Philippines. They began to encroach upon the Mekong-Mamberamo language area, composed of a large number of diverse lineages which had converged structurally over a long period of time due to shared historical links.

When the Austronesians began to settle Flores around 2500-1500 BCE, they encountered speakers of Mekong-Mamberamo type languages who were numerically dominant. The Austronesian settlers enjoyed a relatively high level of prestige over the pre-Austronesian inhabitants, who began to shift from their language to the Austronesian language of the settlers. These Mekong-Mamberamo speakers shifted to Austronesian in a relatively short period of time, leaving a heavy grammatical influence but very little lexical influence. The Austronesian settlers then integrated into the local population and adopted the restructured, Mekong-Mamberamo-like Austronesian language spoken by the majority of the population: Proto-Central Flores (or an immediate ancestor thereof). This ancestral community then differentiated *in situ* into the modern Central Flores languages with no further splits, forming an archetypal linkage (see Elias 2018).

In the scenario just described, the fact that Central Flores languages retain very little lexicon but much of the grammatical structure of the pre-Austronesian languages is explained. However, I do propose that lexical influence from the pre-Austronesian languages of Flores can be seen in one domain: the non-productive Rongga numeral system (see section). This counting system is full of inexplicable oddities: the lower numerals do not resemble any other known languages of the area, the strategy for forming higher numerals is unusual, and several of the numerals contain consonant clusters that violate Rongga phonotactics. All of these point to a now-extinct pre-Austronesian source language whose relatives have all vanished.

Under my historical scenario, I must account for the fact that languages of West Flores and East Flores conform less to the Mekong-Mamberamo typological profile. It is possible that

the pre-Austronesian languages of Flores were themselves diverse, and only the language (or languages) of Central Flores had Mekong-Mamberamo typology to begin with. The Mekong-Mamberamo hypothesis does not claim that all languages across the entire area show all of the features, merely that many of them do. A very likely possibility is that the relevant factors at the time of contact were different in Central Flores and the rest of Flores, such that there was less substratum interference in East and West Flores. In East and West Flores, it is possible that the pre-Austronesian population was less dense, and the ratio of Austronesian settlers to pre-Austronesian inhabitants was higher. This would be expected to lead to less structural interference. Another potential factor is the degree of bilingualism; perhaps the pre-Austronesian and Austronesian populations did not integrate with each other in East and West Flores, and the pre-Austronesians vanished along with their languages rather than shifting to the language of the Austronesian settlers. Another possibility is that the transition took place over longer time periods in East and West Flores, leading to less structural interference in the target language due to more complete bilingualism.

There is another logical possibility which should be mentioned here: it is possible that the level of substratum interference was comparable across all of Flores, and that all Flores languages once resembled the Central Flores languages. Then, morphological complexity was reinvented in the languages of East and West Flores. In this case, we would expect that the morphology seen in East and West Flores would be innovative morphology, not retentions from earlier stages of Austronesian. Overall, it is a mixed picture and the evidence is somewhat inconclusive: much of the morphology, especially in East Flores, seems to be conservative, but there is clearly innovative plural marking in Lamaholot. Similarly, the conjugational systems of Sika and Lamaholot appear at first glance to be conservative, because they are cognate with many of the widespread Central Malayo-Polynesian verbal conjugation systems in the area. However, this may be a hasty judgment, because the verbal conjugation systems are themselves transparently derived from the independent pronouns, and therefore may plausibly have grammaticalised separately in different areas.

Note that in McWhorter's account, the settlers from Sulawesi would have arrived speaking something similar to *Tukang Besi*, which would have become simplified as the (already Austronesian-speaking) inhabitants of Central Flores shifted to the newcomers language. On the reasonable assumption that the settlers from Sulawesi enjoyed high prestige and did not borrow massive amounts of vocabulary from the shifting speakers, we would expect the resulting Central Flores language to resemble *Tukang Besi* lexically, but they do not. If the settlers from Sulawesi did not have high prestige, that raises the question of why the pre-existing population of Central Flores would have shifted to their language at all. The distinct lack of lexical resemblances between Sulawesi and Flores is acknowledged by McWhorter himself: 'lexical and grammatical data in support of this scenario are lacking' (McWhorter 2019; p195). This lack of lexical evidence in support of his theory is a serious problem and must be accounted for.



## 8 Conclusion

The title of this paper poses a question: are the Central Flores languages really typologically unusual? In this paper, I have shown that the Central Flores languages display many of the typological features common in the Mekong-Mamberamo linguistic area proposed by Gil (2015). I argue that the typological profile of the Central Flores languages is best explained by structural interference from a now-extinct substrate language with Mekong-Mamberamo typology. This interference occurred at the time of the original Austronesian settlement of Flores around 2500-1500 BCE.

In contrast, McWhorter (2019) claims that the isolating typology of the Central Flores languages is due to the more general process of simplification that occurs under conditions of language shift, regardless of the typology of the languages involved. He proposes that a group of settlers arrived in Central Flores from Sulawesi in the relatively recent past (once it was already inhabited by Austronesian speakers) and the Central Flores population shifted to their language. This argument suffers from a number of flaws. The most glaring among them are that: 1) it does not account for aspects of the Central Flores language which are not simple, such as their numeral classifier systems, 2) it is undermined by the lack of lexical evidence linking Sulawesi to Flores, and 3) it is less economical to posit two separate contact events (one between Austronesian and non-Austronesian, one between Central Flores and Sulawesi) than a single contact event.

Although there are no non-Austronesian languages spoken on Flores anymore, the non-Austronesian Timor-Alor-Pantar languages are also spoken in the Lesser Sunda Islands. An obvious question is whether or not the substrate language of Central Flores was related to the modern Timor-Alor-Pantar language. This question needs to be investigated further, but the Timor-Alor-Pantar languages are significantly more morphologically complex than the Central Flores languages and are tentatively assigned to the Trans-New Guinea language family based on pronominal evidence. If the Trans-New Guinea status of the Timor-Alor-Pantar languages is confirmed, then it becomes more likely that the Timor-Alor-Pantar languages are a more recent back-migration from the New Guinea mainland to the Lesser Sundas, rather than a relict group of the non-Austronesian speakers who inhabited the area before Austronesian settlement.

In a discipline such as historical linguistics, we cannot hope for experimental replication of results to validate our hypotheses. Instead, we can bolster our historical hypotheses by developing them with one set of data in mind, and showing that they accurately capture the facts in a new set of previously unseen data. In the case of Gil's Mekong-Mamberamo hypothesis, this comes as close to the gold standard as possible: the Mekong-Mamberamo typological features describe the typology of the Central Flores languages very closely, even though the proposal was not developed with the Central Flores languages in mind. Thus, this paper serves not only as a contribution to the prehistory of Central Flores, but also as a practical validation of Gil's Mekong-Mamberamo hypothesis.

## 9 Bibliography

- Aoki, Eriko, and Satoshi Nakagawa. Endenese-English Dictionary. unpublished, 1993.
- Arka, I Wayan. Bahasa Rongga: Deskripsi, Tipologi, Dan Teori [The Rongga Language: Description, Typology and Theory]. Jakarta: Atma Jaya University Press, 2016.
- Arka, I Wayan, Fransiscus Seda, Antonius Gelang, Yohanes Nani, and Ivan Ture. A Rongga-English Dictionary with English-Rongga Finderlist. Jakarta: Atma Jaya University Press, 2011.
- Arndt, Paul. Wörterbuch Der Ngadhasprache. Posieux: Koninklijk Instituut voor Taal-, Land- en Volkenkunde, 1961.
- . Grammatik der Solor-Sprache. Ende: Arnoldus-Druckerei, 1937.
- . Lionesisch-Deutsches Wörterbuch. Ende: Arnoldus-Druckerei, 1933a.
- . Grammatik Der Ngada-Sprache. Bandung: A. C. Nix & Co., 1933b.
- Baird, Louise. A Grammar of Kéo: An Austronesian Language of East Nusantara. Australian National University, 2002.
- Bellwood, Peter. Prehistory of the Indo-Malayan Archipelago. Honolulu: University of Hawaii Press, 1997.
- Blust, Robert. The Austronesian Languages: Revised Edition. Canberra: Asia-Pacific Linguistics, 2013.
- . Is There a Bima-Sumba Subgroup? *Oceanic Linguistics* 47, no. 1 (2008): 45-113.
- Blust, Robert and Stephen Trussel. The Austronesian Comparative Dictionary, Online Edition. Available at <http://www.trussel2.com/ACD/>. 2019.
- Brown, P., T. Sutnika, M. J. Morwood, R. P. Soejono, Jatmiko, E. Wayhu Saptomo, and Rokus Awe Due. A New Small-Bodied Hominin from the Late Pleistocene of Flores, Indonesia. *Nature* 431 (2004): 1055-61.
- Donohue, Mark. The Palue Passive: From Pragmatic Construction to Grammatical Device. In *The Many Faces of Austronesian Voice Systems: Some New Empirical Studies*, edited by I Wayan Arka and Malcolm Ross, 5985. Canberra: Pacific Linguistics, 2005.
- . Bata Lu'a ca (Palu'e-Indonesian-English dictionary). Unpublished, 2003.
- . A Grammar of Tukang Besi. Vol. 20. Mouton Grammar Library. Berlin: De Gruyter Mouton, 1999.
- Elias, Alexander. Lio and the Central Flores languages. Leiden: MA Thesis 2018.
- Erb, Maribeth. The Manggarais. Singapore: Times Edition, 1999.
- Fenk-Oczlon, Gertraud and August Fenk. Complexity trade-offs between the subsystems of language. In *Language Complexity: Typology, contact, change*, edited by Matti Miestamo, Kaius Sinnemäki and Fred Karlsson, 43-65. Vol. 94. *Studies in Language Companion Series*. Amsterdam: John Benjamins, 2008.
- Forth, Gregory. Why the Porcupine is Not a Bird: Explorations in the Folk Zoology of an Eastern Indonesian People. Toronto: University of Toronto Press, 2019.
- . Human beings and other people: Classification of human group and categories among the Nage of Flores (eastern Indonesia).” In *Bijdragen tot de Taal-, Land- en Volkenkunde*

- 165 (2009): 493-514.
- . Images of the wildman in Southeast Asia: An anthropological perspective. London and New York: Routledge, 2008.
  - . Nage birds: Classification and symbolism among an eastern Indonesian people. London and New York: Routledge, 2004.
  - . Beneath the Volcano: Religion, Cosmology and Spirit Classification Among the Nage of Eastern Indonesia. Leiden: Koninklijk Instituut voor Taal-, Land- en Volkenkunde, 1998.
  - Fricke, Hanna. Traces of language contact: The Flores-Lembata languages in eastern Indonesia. Leiden: PhD thesis, 2019.
  - . Topics in the Grammar of Hewa: A Variety of Sika in Eastern Flores, Indonesia. Vol. 45. Languages of the World. Munich: LINCOM, 2014.
  - Gil, David. The Mekong-Mamberamo Linguistic Area. In Languages of Mainland Southeast Asia: The State of the Art, edited by Nick J. Enfield and Bernard Comrie, 649:266355. Pacific Linguistics. Berlin: De Gruyter Mouton, 2015.
  - . Isolating-Monocategorical-Associational Language. In Categorization in Cognitive Science, edited by H. Cohen and C. Lefebvre, 34779. Oxford: Elsevier, 2005.
  - Hull, Geoffrey. The Basic Lexical Affinities of Timors Austronesian Languages: A Preliminary Investigation. Studies in the Languages and Cultures of East Timor 1 (1998): 97-174.
  - Ingram, Arthur. Shepherding Tools and Customs. London: Shire Publications, 1977.
  - Thomason, Sarah and Terrence Kaufman. Language Contact, Creolization, and Genetic Linguistics. Berkeley: University of California Press, 1988.
  - Keraf, Gregorius. Morfologi Dialek Lamalera [Morphology of the Lamalera Dialect]. Jakarta: Universitas Indonesia, PhD dissertation, 1978.
  - Levi, Ferdinandus. A Preliminary Study of Lionese. Sanata Dharma Teachers Training Institute: BA thesis, 1978.
  - Lewis, E. Douglas. The Stanger-kings of Sikka: With an Integrated Edition of Two Manuscripts on the Origin and History of the Rajadom of Sikka. Vol. 257. Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde. Leiden: KITLV.
  - McDonnell, Bradley. A preliminary description of Ende phonology. Journal of the Southeast Asian Linguistics Society 2 (2009): 197-288.
  - McWhorter, John. The Radically Isolating Languages of Flores: A Challenge to Diachronic Theory. Journal of Historical Linguistics 9, no. 2 (2019): 177-207.
  - . Is Radical Analycity Normal?: Implications of Niger-Congo and Southeast Asia for Typology and Diachronic Theory. In Cyclical Change Continued, edited by Elly van Gelderen, 49-92. Amsterdam: John Benjamins, 2016.
  - . Affixless in Austronesian: Why Flores Ia a Puzzle and What to Do about It. In Linguistic Simplicity and Complexity: Why Do Languages Undress?, edited by John McWhorter, 1:223-60. Language Contact and Bilingualism. Berlin: De Gruyter Mouton, 2011.

- . Why Does a Language Undress?: Strange Cases of Indonesia. In *Language Complexity: Typology, Contact, Change*, edited by Matti Miestamo, Kaius Sinnemäki, and Fred Karlsson, 167-90. Amsterdam: John Benjamins, 2008.
- . *Language Interrupted: Signs of Non-Native Acquisition in Standard Language Grammars*. Oxford: Oxford University Press, 2007.
- . The Worlds Simplest Grammars Are Creole Grammars. *Linguistic Typology* 5, no. 2 (2001): 125-66.
- Nagaya, Naonori. *The Lamaholot Language of Eastern Indonesia*. Rice University, 2011.
- Nishiyama, Kunio and Herman Kelen. *A Grammar of Lamaholot, Eastern Indonesia: The Morphology and Syntax of the Lewoingu Dialect [Languages of the World/Materials 467]*. Munich: LINCOM Europa, 2007.
- Pampus, Karl Heinz. *Koda Kiwa: Dreisprachiges Wörterbuch des Lamaholot (Dialekt von Lewolema)*. Stuttgart: Steiner, 1999.
- Pareira, Mandalangi and E. Douglas Lewis. *Kamus Sara Sikka-Bahasa Indonesia [Sika-Indonesian Dictionary]*. Ende: Nusa Indah, 1998.
- Schapper, Antoinette, and Marian Klamer. Numeral Systems in the Alor-Pantar Languages. In *The Alor-Pantar Languages: History and Typology*, edited by Marian Klamer, 285-336. Berlin: Language Science Press, 2014.
- Semiun, Agustinus. Clitics of Manggarai Language: The Case of Central Manggarai Dialect in West Flores Indonesia. *International Journal of Development Research* 7, no. 1 (2017): 11380-11387.
- Thomason, Sarah, and Terrence Kaufman. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press, 1988.
- Thurston, William. *Processes of Change in the Languages of Northwestern New Britain*. Canberra: Pacific Linguistics, 1987.
- Van Bakkum, W. Warloka-Todo-Pongkor: Een Brok Geschiedenis van Manggarai (West-Flores). *Cultureel Indië* 6 (1944): 144-52.
- Verheijen, Jilis. *Komodo: Het eiland, het volk en de taal*. Leiden: Martinus Nijhoff, 1982.
- . *Kamus Manggarai [Manggarai Dictionary]*. The Hague: Koninklijk Instituut voor Taal-, Land- en Volkenkunde, 1967.